

# New Strong Converse for Asymmetric Broadcast Channels

Yasutada Oohama

University of Electro-Communications, Tokyo, Japan

Email: oohama@uec.ac.jp

**Abstract**—We consider the discrete memoryless asymmetric broadcast channels. We prove that the error probability of decoding tends to one exponentially for rates outside the capacity region and derive an explicit lower bound of this exponent function. We shall demonstrate that the information spectrum approach is quite useful for investigating this problem.

**Keywords**—Discrete memoryless channels, asymmetric broadcast channels, strong converse theorem, exponent of correct probability of decoding

## I. ASYMMETRIC BROADCAST CHANNELS

Let  $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$  be finite sets. The broadcast channel we study in this paper is defined by a discrete memoryless channel specified with the following stochastic matrix:

$$W \triangleq \{W(y, z|x)\}_{(x,y,z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}}. \quad (1)$$

Here the set  $\mathcal{X}$  corresponds to a channel input and the sets  $\mathcal{Y}$  and  $\mathcal{Z}$  correspond to two channel outputs. Let  $X^n$  be a random variable taking values in  $\mathcal{X}^n$ . We write an element of  $\mathcal{X}^n$  as  $x^n = x_1 x_2 \cdots x_n$ . Suppose that  $X^n$  has a probability distribution on  $\mathcal{X}^n$  denoted by  $p_{X^n} = \{p_{X^n}(x^n)\}_{x^n \in \mathcal{X}^n}$ . Similar notations are adopted for other random variables. Let  $Y^n \in \mathcal{Y}^n$  and  $Z^n \in \mathcal{Z}^n$  be random variables obtained as the channel output by connecting  $X^n$  to the input of channel. We write a conditional distribution of  $(Y^n, Z^n)$  on given  $X^n$  as

$$W^n = \{W^n(y^n, z^n|x^n)\}_{(x^n, y^n, z^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n}.$$

Since the channel is memoryless, we have

$$W^n(y^n, z^n|x^n) = \prod_{t=1}^n W(y_t, z_t|x_t). \quad (2)$$

In this paper we deal with the case where the components  $W(z, y|x)$  of  $W$  satisfy the following conditions:

$$W(y, z|x) = W_1(y|x)W_2(z|x). \quad (3)$$

In this case the broadcast channel(BC) is specified with  $(W_1, W_2)$ . Under the assumption of (3), the conditional probability of (2) is given by

$$\begin{aligned} W^n(y^n, z^n|x^n) &= W_1^n(y^n|x^n)W_2^n(z^n|x^n) \\ &= \prod_{t=1}^n W_1(y_t|x_t)W_2(z_t|x_t). \end{aligned}$$

Transmission of messages via the BC is shown in Fig. 1. Let  $K_n$  and  $L_n$  be uniformly distributed random variables taking values in message sets  $\mathcal{K}_n$  and  $\mathcal{L}_n$ , respectively. The random variable  $K_n$  is a message sent to the receiver 1. The random

variable  $L_n$  is a message sent to the receivers 1 and 2. A sender transforms  $K_n$  and  $L_n$  into a transmitted sequence  $X^n$  using an encoder function  $\varphi^{(n)}$  and sends it to the receivers 1 and 2. In this paper we assume that the encoder function  $\varphi^{(n)}$  is a stochastic encoder. In this case,  $\varphi^{(n)}$  is a stochastic matrix given by

$$\varphi^{(n)} = \{\varphi^{(n)}(x^n|k, l)\}_{(k,l,x^n) \in \mathcal{K}_n \times \mathcal{L}_n \times \mathcal{X}^n},$$

where  $\varphi^{(n)}(x^n|k, l)$  is a conditional probability of  $x^n \in \mathcal{X}^n$  given message pair  $(k, l) \in \mathcal{K}_n \times \mathcal{L}_n$ . The joint probability mass function on  $\mathcal{K}_n \times \mathcal{L}_n \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$  is given by

$$\begin{aligned} \Pr\{(K_n, L_n, X^n, Y^n, Z^n) = (k, l, x^n, y^n, z^n)\} \\ = \frac{\varphi^{(n)}(x^n|k, l)}{|\mathcal{K}_n||\mathcal{L}_n|} \prod_{t=1}^n W_1(y_t|x_t) W_2(z_t|y_t), \end{aligned}$$

where  $|\mathcal{K}_n|$  is a cardinality of the set  $\mathcal{K}_n$ . The decoding functions at the receiver 1 and the receiver 2, respectively, are denoted by  $\psi_1^{(n)}$  and  $\psi_2^{(n)}$ . Those functions are formally defined by

$$\psi_1^{(n)} : \mathcal{Y}^n \rightarrow \mathcal{K}_n \times \mathcal{L}_n, \psi_2^{(n)} : \mathcal{Z}^n \rightarrow \mathcal{L}_n.$$

The average error probabilities of decoding at the receivers 1 and 2 are defined by

$$\begin{aligned} P_{e,1}^{(n)} &= P_e^{(n)}(\varphi^{(n)}, \psi_1^{(n)}) \triangleq \Pr\{\psi_1^{(n)}(Y^n) \neq (K_n, L_n)\}, \\ P_{e,2}^{(n)} &= P_e^{(n)}(\varphi^{(n)}, \psi_2^{(n)}) \triangleq \Pr\{\psi_2^{(n)}(Z^n) \neq L_n\}. \end{aligned}$$

Furthermore, we set

$$\begin{aligned} P_e^{(n)} &= P_e^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}) \\ &\triangleq \Pr\{\psi_1^{(n)}(Y^n) \neq (K_n, L_n) \text{ or } \psi_2^{(n)}(Z^n) \neq L_n\} \end{aligned}$$

It is obvious that we have the following relation.

$$P_e^{(n)} \leq P_{e,1}^{(n)} + P_{e,2}^{(n)}. \quad (4)$$

For  $k \in \mathcal{K}_n$  and  $l \in \mathcal{L}_n$ , set

$$\begin{aligned} \mathcal{D}_1(k, l) &\triangleq \{y^n : \psi_1^{(n)}(y^n) = (k, l)\}, \\ \mathcal{D}_2(l) &\triangleq \{z^n : \psi_2^{(n)}(z^n) = l\}. \end{aligned}$$

The families of sets  $\{\mathcal{D}_1(k, l)\}_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n}$  and  $\{\mathcal{D}_2(l)\}_{l \in \mathcal{L}_n}$  are called the decoding regions. Using the decoding region,  $P_e^{(n)}$  can be written as

$$\begin{aligned} P_e^{(n)} &= \frac{1}{|\mathcal{K}_n||\mathcal{L}_n|} \sum_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n, z^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n: \\ y^n \in \mathcal{D}_1^c(k, l) \text{ or } z^n \in \mathcal{D}_2^c(l)}} 1 \\ &\quad \times \varphi^{(n)}(x^n|k, l) W_1^n(y^n|x^n) W_2^n(z^n|x^n). \end{aligned}$$

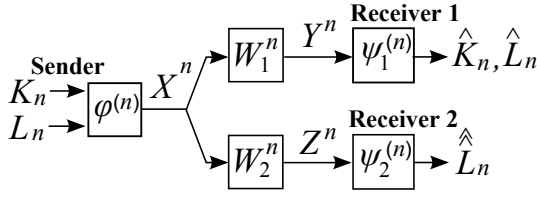


Fig. 1. Transmission of messages via the BC.

Set

$$\begin{aligned} P_c^{(n)} &= P_c^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}) \\ &\triangleq 1 - P_e^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}). \end{aligned}$$

The quantity  $P_c^{(n)}$  is called the average correct probability of decoding. This quantity has the following form

$$\begin{aligned} P_c^{(n)} &= \frac{1}{|\mathcal{K}_n| |\mathcal{L}_n|} \sum_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n, z^n) \in \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n: \\ y^n \in \mathcal{D}_1(k,l), z^n \in \mathcal{D}_2(l)}} 1 \\ &\quad \times \varphi^{(n)}(x^n|k, l) W_1^n(y^n|x^n) W_2^n(z^n|x^n). \end{aligned}$$

For a given  $(\varepsilon_1, \varepsilon_2) \in (0, 1)^2$ , a rate pair  $(R_1, R_2)$  is  $(\varepsilon_1, \varepsilon_2)$ -achievable if there exists a sequence of triples  $\{(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)})\}_{n=1}^\infty$  such that for any  $\delta > 0$  and for any  $n$  with  $n \geq n_0 = n_0(\varepsilon_1, \varepsilon_2, \delta)$

$$\begin{aligned} P_{e,i}^{(n)}(\varphi^{(n)}, \psi_i^{(n)}) &\leq \varepsilon_i, i = 1, 2, \\ \frac{1}{n} \log |\mathcal{K}_n| &\geq R_1 - \delta, \frac{1}{n} \log |\mathcal{L}_n| \geq R_2 - \delta. \end{aligned}$$

The set that consists of all  $(\varepsilon_1, \varepsilon_2)$ -achievable rate pair is denoted by  $\mathcal{C}_{ABC}(\varepsilon_1, \varepsilon_2|W_1, W_2)$ . For a given  $\varepsilon \in (0, 1)$ , a pair  $(R_1, R_2)$  is  $\varepsilon$ -achievable if there exists a sequence of triples  $\{(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)})\}_{n=1}^\infty$  such that for any  $\delta > 0$  and for any  $n$  with  $n \geq n_0 = n_0(\varepsilon, \delta)$

$$\begin{aligned} P_e^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}) &\leq \varepsilon, \\ \frac{1}{n} \log |\mathcal{K}_n| &\geq R_1 - \delta, \frac{1}{n} \log |\mathcal{L}_n| \geq R_2 - \delta. \end{aligned}$$

The set that consists of all  $\varepsilon$ -achievable rate pair is denoted by  $\mathcal{C}_{ABC}(\varepsilon|W_1, W_2)$ . It is obvious that for  $0 \leq \varepsilon_1 + \varepsilon_2 \leq 1$ , we have

$$\mathcal{C}_{ABC}(\varepsilon_1, \varepsilon_2|W_1, W_2) \subseteq \mathcal{C}_{ABC}(\varepsilon_1 + \varepsilon_2|W_1, W_2).$$

We set

$$\mathcal{C}_{ABC}(W_1, W_2) \triangleq \bigcap_{\varepsilon \in (0,1)} \mathcal{C}_{ABC}(\varepsilon|W_1, W_2),$$

which is called the capacity region of the ABC. The two maximum error probabilities of decoding are defined by as follows:

$$\begin{aligned} P_{e,m,1}^{(n)} &= P_{e,m,1}^{(n)}(\varphi^{(n)}, \psi_1^{(n)}) \\ &\triangleq \max_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \Pr\{\psi_1^{(n)}(Y^n) \neq k | K_n = k, L_n = l\}, \\ P_{e,m,2}^{(n)} &= P_{e,m,2}^{(n)}(\varphi^{(n)}, \psi_2^{(n)}) \\ &\triangleq \max_{l \in \mathcal{L}_n} \Pr\{\psi_2^{(n)}(Z^n) \neq l | L_n = l\}. \end{aligned}$$

Based on those quantities, we define the maximum capacity region as follows. For a given  $(\varepsilon_1, \varepsilon_2) \in (0, 1)^2$ , a pair  $(R_1, R_2)$  is  $(\varepsilon_1, \varepsilon_2)$ -achievable if there exists a sequence of triples  $\{(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)})\}_{n=1}^\infty$  such that for any  $\delta > 0$  and for any  $n$  with  $n \geq n_0 = n_0(\varepsilon_1, \varepsilon_2, \delta)$

$$\begin{aligned} P_{e,m,i}^{(n)}(\varphi^{(n)}, \psi_i^{(n)}) &\leq \varepsilon_i, i = 1, 2, \\ \frac{1}{n} \log |\mathcal{K}_n| &\geq R_1 - \delta, \frac{1}{n} \log |\mathcal{L}_n| \geq R_2 - \delta. \end{aligned}$$

The set that consists of all  $(\varepsilon_1, \varepsilon_2)$ -achievable rate pair is denoted by  $\mathcal{C}_{m,ABC}(\varepsilon_1, \varepsilon_2|W_1, W_2)$ . We set

$$\mathcal{C}_{m,ABC}(W_1, W_2) = \bigcap_{(\varepsilon_1, \varepsilon_2) \in (0,1)^2} \mathcal{C}_{m,ABC}(\varepsilon_1, \varepsilon_2|W_1, W_2),$$

which is called the maximum capacity region of the ABC. It is obvious that

$$\mathcal{C}_{m,ABC}(\varepsilon_1, \varepsilon_2|W_1, W_2) \subseteq \mathcal{C}_{ABC}(\varepsilon_1, \varepsilon_2|W_1, W_2).$$

To describe previous works on  $\mathcal{C}_{m,ABC}(\varepsilon_1, \varepsilon_2|W_1, W_2)$  and  $\mathcal{C}_{ABC}(\varepsilon_1, \varepsilon_2|W_1, W_2)$ , we introduce an auxiliary random variable  $U$  taking values in a finite set  $\mathcal{U}$ . We assume that the joint distribution of  $(U, X, Y, Z)$  is

$$p_{UXYZ}(u, x, y, z) = p_U(u) p_{X|U}(x|u) W_1(y|x) W_2(z|x).$$

The above condition is equivalent to  $Y \leftrightarrow X \leftrightarrow Z$  and  $U \leftrightarrow X \leftrightarrow (Y, Z)$ . Define the set of probability distribution  $p = p_{UXYZ}$  of  $(U, X, Y, Z) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$  by

$$\begin{aligned} \mathcal{P}(W_1, W_2) &\triangleq \{p : |\mathcal{U}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}| + |\mathcal{Z}|\} + 1, \\ p_{Y|X} &= W_1, p_{Z|X} = W_2, \\ U &\leftrightarrow X \leftrightarrow (Y, Z), Y \leftrightarrow X \leftrightarrow Z\}. \end{aligned}$$

Set

$$\begin{aligned} \mathcal{C}(p) &\triangleq \{(R_1, R_2) : R_1, R_2 \geq 0, \\ &\quad R_1 \leq I_p(X; Y|U), \\ &\quad R_2 \leq I_p(U; Z), \\ &\quad R_1 + R_2 \leq I_p(X; Y)\}, \\ \mathcal{C}(W_1, W_2) &= \bigcup_{p \in \mathcal{P}(W_1, W_2)} \mathcal{C}(p). \end{aligned}$$

We can show that the above functions and sets satisfy the following property.

*Property 1:*

a) Set

$$\begin{aligned} \mathcal{C}_{\text{ext}}(p) &\triangleq \{(r_1, r_2, r_3) : r_1, r_2, r_3 \geq 0, \\ &\quad r_1 \leq I_p(X; Y|U), \\ &\quad r_2 \leq I_p(U; Z), \\ &\quad r_3 \leq I_p(X; Y)\}, \\ \mathcal{C}_{\text{ext}}(W_1, W_2) &= \bigcup_{p \in \mathcal{P}(W_1, W_2)} \mathcal{C}_{\text{ext}}(p). \end{aligned}$$

Then we have

$$\begin{aligned} \mathcal{C}(W_1, W_2) &= \{(R_1, R_2) : (R_1, R_2, R_1 + R_2) \\ &\quad \in \mathcal{C}_{\text{ext}}(W_1, W_2)\} \end{aligned}$$

- b) The cardinality bound in  $\mathcal{P}(W_1, W_2)$  is sufficient to describe  $\mathcal{C}(W_1, W_2)$  and  $\mathcal{C}_{\text{ext}}(W_1, W_2)$ .
- c) The region  $\mathcal{C}(W_1, W_2)$  is a closed convex subset of  $\mathbb{R}_+^2$  and the region  $\mathcal{C}_{\text{ext}}(W_1, W_2)$  is a closed convex subset of  $\mathbb{R}_+^4$ , where

$$\mathbb{R}_+^2 \triangleq \{(R_1, R_2) : R_1 \geq 0, R_2 \geq 0\},$$

$$\mathbb{R}_+^3 \triangleq \{(r_1, r_2, r_3) : r_i \geq 0, i = 1, 2, 3\}.$$

- d) The region  $\mathcal{C}(W_1, W_2)$  is always contained in, and may coincide with, the triangle with vertices  $(0, 0)$ ,  $(C(W_1), 0)$ ,  $(0, C(W_1))$ , where  $C(W_1)$  is the capacity of the channel  $W_1$ . The point  $(C(W_1), 0)$  always belongs to  $\mathcal{C}(W_1, W_2)$ . In general, the upper boundary of  $\mathcal{C}(W_1, W_2)$  contains a line segment of slope  $-1$  going through the point  $(C(W_1), 0)$  but this line segment may reduce to the point  $(C(W_1), 0)$ .

Property 1 part a) is obvious. Property 1 parts b), c) are well known result. Property 1 part d) is found in [2]. Proof of those properties are omitted.

The broadcast channel was posed and investigated by Cover [1]. The capacity region of the ABC was given by Körner and Marton [2]. They called the ABC the broadcast channels with degraded message sets. Körner and Marton [2] obtained the following result.

*Theorem 1 (Körner and Marton [2]):* For each fixed  $\varepsilon \in (0, 1)$  and for any  $(W_1, W_2)$ , we have

$$\mathcal{C}_{\text{m,ABC}}(\varepsilon, \varepsilon | W_1, W_2) = \mathcal{C}_{\text{ABC}}(W_1, W_2) \\ = \mathcal{C}(W_1, W_2).$$

To prove this theorem they used a combinatorial lemma called “the blowing up lemma”. Their method used to prove the above theorem was extended to the method called the entropy and image size characterization by Csiszár and Körner [3], where they obtained the following result.

*Theorem 2 (Csiszár and Körner [3]):* For  $\varepsilon \in (0, 1/2)$ ,

$$\mathcal{C}_{\text{ABC}}(\varepsilon, \varepsilon | W_1, W_2) = \mathcal{C}_{\text{ABC}}(W_1, W_2).$$

Csiszár and Körner [3] applied the method of entropy and image size characterization to other coding problems in multi-user information theory to prove strong converse theorems for those problems. Universally attainable error exponents for rates inside the capacity region  $\mathcal{C}(W_1, W_2)$  was studied by Körner and Sgarro [4] and Kaspri and Merhav [5].

To examine an asymptotic behavior of  $P_c^{(n)}$  for rates outside the capacity region  $\mathcal{C}(W_1, W_2)$ , we define the following quantity.

$$G^{(n)}(R_1, R_2 | W_1, W_2) \\ \triangleq \min_{\substack{(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}) : \\ (1/n) \log |\mathcal{K}_n| \geq R_1, \\ (1/n) \log |\mathcal{L}_n| \geq R_2}} \left( -\frac{1}{n} \right) \log P_c^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}).$$

By time sharing we have that

$$G^{(n+m)} \left( \frac{nR_1 + mR'_1}{n+m}, \frac{nR_2 + mR'_2}{n+m} \middle| W_1, W_2 \right) \\ \leq \frac{nG^{(n)}(R_1, R_2 | W_1, W_2) + mG^{(m)}(R'_1, R'_2 | W_1, W_2)}{n+m}. \quad (5)$$

Choosing  $R_1 = R'_1$  and  $R_2 = R'_2$  in (5), we obtain the following subadditivity property on  $\{G^{(n)}(R_1, R_2 | W_1, W_2)\}_{n \geq 1}$ :

$$G^{(n+m)}(R_1, R_2 | W_1, W_2) \\ \leq \frac{nG^{(n)}(R_1, R_2 | W_1, W_2) + mG^{(m)}(R_1, R_2 | W_1, W_2)}{n+m},$$

from which we have that  $G(R_1, R_2 | W_1, W_2)$  exists and satisfies the following:

$$\lim_{n \rightarrow \infty} G^{(n)}(R_1, R_2 | W_1, W_2) = \inf_{n \geq 1} G^{(n)}(R_1, R_2 | W_1, W_2).$$

Set

$$G(R_1, R_2 | W_1, W_2) \triangleq \lim_{n \rightarrow \infty} G^{(n)}(R_1, R_2 | W_1, W_2), \\ \mathcal{G}(p_{XY}) \triangleq \{(R_1, R_2, G) : G \geq G(R_1, R_2 | W_1, W_2)\}.$$

The exponent function  $G(R_1, R_2 | W_1, W_2)$  is a convex function of  $(R_1, R_2)$ . In fact, from (5), we have that for any  $\alpha \in [0, 1]$

$$G(\alpha R_1 + \bar{\alpha} R'_1, \alpha R_2 + \bar{\alpha} R'_2 | W_1, W_2) \\ \leq \alpha G(R_1, R_2 | W_1, W_2) + \bar{\alpha} G(R'_1, R'_2 | W_1, W_2).$$

The region  $\mathcal{R}(W_1, W_2)$  is also a closed convex set. Our main aim is to find an explicit characterization of  $\mathcal{R}(W_1, W_2)$ . In this paper we derive an explicit outer bound of  $\mathcal{R}(W_1, W_2)$  whose section by the plane  $G = 0$  coincides with  $\mathcal{C}(W_1, W_2)$ .

## II. MAIN RESULT

In this section we state our main result. Before describing the result, we state that the region  $\mathcal{C}_{\text{ext}}(W_1, W_2)$  can be expressed with two families of supporting hyperplanes. We define the set of probability distribution  $p = p_{UXYZ}$  of  $(U, X, Y, Z) \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$  by

$$\mathcal{P}_{\text{sh}}(W_1, W_2) \\ \triangleq \{p = p_{UXYZ} : |\mathcal{U}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}| + |\mathcal{Z}| - 1\}, \\ p_{Y|X} = W_1, p_{Z|Y} = W_2, \\ Y \leftrightarrow X \leftrightarrow Z, U \leftrightarrow X \leftrightarrow (Y, Z)\}, \\ \mathcal{Q} \triangleq \{q = q_{UXYZ} : |\mathcal{U}| \leq |\mathcal{Y}| + |\mathcal{Z}| - 1\}.$$

We set

$$\begin{aligned}
& C^{(\gamma, \mu)}(W_1, W_2) \\
& \triangleq \max_{p \in \mathcal{P}_{\text{sh}}(W_1, W_2)} \{ \gamma \mu I_p(X; Y|U) + \gamma \bar{\mu} I_p(U; Z) \\
& \quad + \bar{\gamma} I_p(X; Y) \}, \\
& \tilde{C}^{(\alpha, \beta, \gamma, \mu)}(W_1, W_2) \\
& \triangleq \max_{q \in \mathcal{Q}} \{ -\alpha D(q_Y|_{XZU} || W_1|_{q_{XZU}}) \\
& \quad - \beta D(q_Z|_{XYU} || W_2|_{q_{XYU}}) \\
& \quad + \gamma \mu I_q(X; Y|U) + \gamma \bar{\mu} I_q(U; Z) \\
& \quad + \bar{\gamma} I_q(X; Y) \}, \\
& \mathcal{C}_{\text{ext, sh}}(W_1, W_2) \\
& = \bigcap_{\substack{\gamma \in [0, 1], \\ \mu \in [0, 1/2]}} \{ (r_1, r_2, r_3) : \\
& \quad \gamma[\mu r_1 + \bar{\mu} r_2] + \bar{\gamma} r_3 \\
& \quad \leq C^{(\gamma, \mu)}(W_1, W_2) \}, \\
& \tilde{\mathcal{C}}_{\text{ext, sh}}(W_1, W_2) \\
& = \bigcap_{\substack{\alpha, \beta > 0, \\ \gamma \in [0, 1], \\ \mu \in [0, 1/2]}} \{ (r_1, r_2, r_3) : \\
& \quad \gamma[\mu r_1 + \bar{\mu} r_2] + \bar{\gamma} r_3 \\
& \quad \leq \tilde{C}^{(\alpha, \beta, \gamma, \mu)}(W_1, W_2) \}.
\end{aligned}$$

Then we have the following Property.

*Property 2:*

a) The cardinality bound in  $\mathcal{P}_{\text{sh}}(W_1, W_2)$  is sufficient to describe  $\mathcal{C}_{\text{ext, sh}}(W_1, W_2)$ .

b)

$$\mathcal{C}_{\text{ext, sh}}(W_1, W_2) = \tilde{\mathcal{C}}_{\text{ext, sh}}(W_1, W_2) = \mathcal{C}_{\text{ext}}(W_1, W_2).$$

Proof of Property 2 part a) is given in Appendix A. Proof of Property 2 part b) is given in Appendix B. Define

$$\begin{aligned}
& \omega_q^{(\alpha, \beta, \gamma, \mu)}(x, y, z|u) \\
& \triangleq \alpha \log \frac{W_1(y|x)}{q_{Y|XZU}(y|x, z, u)} + \beta \log \frac{W_2(z|x)}{q_{Z|XYU}(z|x, y, u)} \\
& \quad + \gamma \left[ \mu \log \frac{W_1(y|x)}{q_{Y|U}(y|u)} + \bar{\mu} \log \frac{q_{Z|U}(z|u)}{q_Z(z)} \right] \\
& \quad + \bar{\gamma} \log \frac{W_1(y|x)}{q_Y(y)}.
\end{aligned}$$

Furthermore, define

$$\begin{aligned}
& \Lambda^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2) \\
& \triangleq \sum_{\substack{(u, x, y, z) \\ \in \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}}} q_{UXYZ}(u, x, y, z) \\
& \quad \times \exp \left\{ \lambda \omega_q^{(\alpha, \beta, \gamma, \mu)}(x, y, z|u) \right\}, \\
& \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2) \triangleq \log \Lambda^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2), \\
& \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2) \\
& \triangleq \max_q \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2), \\
& F_{\text{ext}}^{(\alpha, \beta, \gamma, \mu, \lambda)}(r_1, r_2, r_3|W_1, W_2) \\
& \triangleq \{ 1 + \lambda[1 + \alpha + \beta + (2 - 3\mu)\gamma] \}^{-1} \\
& \quad \times \left[ \lambda \{ \gamma(\mu r_1 + \bar{\mu} r_2) + \bar{\gamma} r_3 \} - \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2) \right],
\end{aligned}$$

$$\begin{aligned}
& F^{(\alpha, \beta, \gamma, \mu, \lambda)}(R_1, R_2|W_1, W_2) \\
& \triangleq F_{\text{ext}}^{(\alpha, \beta, \gamma, \mu, \lambda)}(R_1, R_2, R_1 + R_2|W_1, W_2) \\
& = \frac{\lambda [(\gamma\mu + \bar{\gamma})R_1 + (\gamma\bar{\mu} + \bar{\gamma})R_2] - \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)}{1 + \lambda[1 + \alpha + \beta + (2 - 3\mu)\gamma]}, \\
& F_{\text{ext}}(r_1, r_2, r_3|W_1, W_2) \\
& \triangleq \sup_{\substack{\alpha, \beta, \lambda > 0, \\ (\gamma, \mu) \in [0, 1] \times [0, 1/2]}} F_{\text{ext}}^{(\alpha, \beta, \gamma, \mu, \lambda)}(r_1, r_2, r_3|W_1, W_2), \\
& F(R_1, R_2|W_1, W_2) \\
& \triangleq F_{\text{ext}}(R_1, R_2, R_1 + R_2|W_1, W_2) \\
& = \sup_{\substack{\alpha, \beta, \lambda > 0, \\ (\gamma, \mu) \in [0, 1] \times [0, 1/2]}} F^{(\alpha, \beta, \gamma, \mu, \lambda)}(R_1, R_2|W_1, W_2), \\
& \overline{\mathcal{R}}(W_1, W_2) \triangleq \{ (R_1, R_2, G) : G \geq F(R_1, R_2|W_1, W_2) \}.
\end{aligned}$$

We can show that the above functions and sets satisfy the following property.

*Property 3:*

a)  $\Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2)$  is a convex function of  $\lambda > 0$ .

b) For every  $q \in \mathcal{Q}$ , we have

$$\begin{aligned}
& \lim_{\lambda \rightarrow +0} \frac{\Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2)}{\lambda} \\
& = -\alpha D(q_Y|_{XZU} || W_1|_{q_{XZU}}) \\
& \quad - \beta D(q_Z|_{XYU} || W_2|_{q_{XYU}}) \\
& \quad - \gamma \mu D(q_Y|_{XU} || W_1|_{q_{XU}}) \\
& \quad + \gamma [\mu I_q(X; Y|U) + \bar{\mu} I_q(U; Z)] + \bar{\gamma} I_q(X; Y).
\end{aligned}$$

c) If  $(r_1, r_2, r_3) \notin \mathcal{C}_{\text{ext}}(W_1, W_2)$ , then we have

$$F_{\text{ext}}(r_1, r_2, r_3|W_1, W_2) > 0.$$

In particular, if  $(R_1, R_2, R_1 + R_2) \notin \mathcal{C}_{\text{ext}}(W_1, W_2)$ , then we have

$$F_{\text{ext}}(R_1, R_2, R_1 + R_2|W_1, W_2) > 0,$$

which implies that if  $(R_1, R_2) \notin \mathcal{C}(W_1, W_2)$ , then we have

$$F(R_1, R_2|W_1, W_2) > 0.$$

Proof of Property 3 is given in Appendix C. Our main result is the following.

*Theorem 3:* For any BC  $(W_1, W_2)$ , we have

$$G(R_1, R_2|W_1, W_2) \geq F(R_1, R_2|W_1, W_2), \quad (6)$$

$$\mathcal{R}(W_1, W_2) \subseteq \overline{\mathcal{R}}(W_1, W_2). \quad (7)$$

It follows from Theorem 3 and Property 3 part c) that if  $(R_1, R_2)$  is outside the capacity region, then the error probability of decoding goes to one exponentially and its exponent is not below  $F(R_1, R_2|W_1, W_2)$ .

From this theorem we immediately obtain the following corollary, which recovers the strong converse theorem by Csiszár and Körner [3].

*Corollary 1:* For each fixed  $\varepsilon \in (0, 1)$ , and for any  $(W_1, W_2)$ , we have

$$\mathcal{C}_{\text{ABC}}(\varepsilon|W_1, W_2) = \mathcal{C}_{\text{ABC}}(W_1, W_2) = \mathcal{C}(W_1, W_2).$$

For each fixed  $\varepsilon \in (0, 1/2)$ , and for any  $(W_1, W_2)$ , we have

$$\begin{aligned} \mathcal{C}_{\text{ABC}}(\varepsilon, \varepsilon | W_1, W_2) &= \mathcal{C}_{\text{ABC}}(W_1, W_2) \\ &= \mathcal{C}(W_1, W_2). \end{aligned}$$

Proof of Theorem 3 will be given in the next section. The strong converse theorem for the single-user channel is proved by a combinatorial method called the method of types, which is developed by Csiszár and Körner [3]. This method is not useful for proving Theorem 3. In fact when we use the method of type, it is very hard to extract a condition related to the Markov chain condition  $U \leftrightarrow X \leftrightarrow Y$ , which the auxiliary random variable  $U \in \mathcal{U}$  must satisfy when  $(R_1, R_2)$  is on the boundary of  $\mathcal{C}(W_1, W_2)$ . Some novel techniques based on the information spectrum method introduced by Han [6] are necessary to prove this theorem.

### III. PROOFS OF THE MAIN RESULTS

We first prove the following lemma.

*Lemma 1:* For any  $\eta > 0$  and for any  $(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)})$  satisfying

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R_1, \frac{1}{n} \log |\mathcal{L}_n| \geq R_2,$$

we have

$$\begin{aligned} P_c^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}) &\leq p_{L_n X^n Y^n Z^n} \left\{ \right. \\ 0 &\leq \frac{1}{n} \log \frac{W_1^n(Y^n | X^n)}{Q_{Y^n | X^n Z^n L_n}^{(i)}(Y^n | X^n, Z^n, L_n)} + \eta, \quad (8) \end{aligned}$$

$$0 \leq \frac{1}{n} \log \frac{W_2^n(Z^n | X^n)}{Q_{Z^n | X^n Y^n L_n}^{(ii)}(Z^n | X^n, Y^n, L_n)} + \eta, \quad (9)$$

$$R_1 \leq \frac{1}{n} \log \frac{W_1^n(Y^n | X^n)}{Q_{Y^n | L_n}^{(iii)}(Y^n | L_n)} + \eta, \quad (10)$$

$$R_2 \leq \frac{1}{n} \log \frac{p_{Z^n | L_n}(Z^n | L_n)}{Q_{Z^n}^{(iv)}(Z^n)} + \eta, \quad (11)$$

$$R_1 + R_2 \leq \frac{1}{n} \log \frac{W_1^n(Y^n | X^n)}{Q_{Y^n}^{(v)}(Y^n)} + \eta \Big\} + 5e^{-n\eta}. \quad (12)$$

In (8), we can choose any conditional distribution  $Q_{Y^n | X^n Z^n L_n}^{(i)}$  on  $\mathcal{Y}^n$  given  $(X^n, Z^n, L_n)$ . In (9), we can choose any conditional distribution  $Q_{Z^n | X^n Y^n L_n}^{(ii)}$  on  $\mathcal{Z}^n$  given  $(X^n, Y^n, L_n)$ . In (10), we can choose any conditional distribution  $Q_{Y^n | L_n}^{(iii)}$  on  $\mathcal{Y}^n$  given  $L_n$ . In (11), we can choose any distribution  $Q_{Z^n}^{(iv)}$  on  $\mathcal{Z}^n$ . In (12), we can choose any distribution  $Q_{Y^n}^{(v)}$  on  $\mathcal{Y}^n$ .

Proof of this lemma is given in Appendix D. For  $t = 1, 2, \dots, n$ , set

$$\begin{aligned} \mathcal{U}_t &\triangleq \mathcal{L}_n \times \mathcal{Y}^{t-1} \times \mathcal{Z}_{t+1}^n, U_t \triangleq (L_n, Y^{t-1}, Z_{t+1}^n) \in \mathcal{U}_t, \\ u_t &\triangleq (l, y^{t-1}, z_{t+1}^n) \in \mathcal{U}_t, \\ \mathcal{V}_t &\triangleq \mathcal{L}_n \times \mathcal{Z}_{t+1}^n, V_t \triangleq (L_n, Z_{t+1}^n) \in \mathcal{V}_t. \end{aligned}$$

For each  $t = 1, 2, \dots, n$ , let  $\kappa_t$  be a natural projection from  $\mathcal{U}_t$  onto  $\mathcal{V}_t$ . Using  $\kappa_t$ , we have  $V_t = \kappa_t(U_t)$ ,  $t = 1, 2, \dots, n$ . From Lemma 1 we have the following.

*Lemma 2:* For any  $\eta > 0$  and for any  $(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)})$  satisfying

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R_1, \frac{1}{n} \log |\mathcal{L}_n| \geq R_2,$$

we have

$$\begin{aligned} P_c^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}) &\leq p_{L_n X^n Y^n Z^n} \left\{ \right. \\ 0 &\leq \frac{1}{n} \sum_{t=1}^n \log \frac{W_1(Y_t | X_t)}{Q_{Y_t | X_t Z_t U_t}^{(i)}(Y_t | X_t Z_t U_t)} + \eta, \\ 0 &\leq \frac{1}{n} \sum_{t=1}^n \log \frac{W_2(Z_t | X_t)}{Q_{Z_t | X_t Y_t U_t}^{(ii)}(Z_t | X_t Y_t U_t)} + \eta, \\ R_1 &\leq \frac{1}{n} \sum_{t=1}^n \log \frac{W_1(Y_t | X_t)}{Q_{Y_t | U_t}^{(iii)}(Y_t | U_t)} \\ &\quad + \frac{1}{n} \sum_{t=1}^n \log \frac{Q_{Z_t | U_t}^{(iii)}(Z_t | U_t)}{Q_{Z_t | V_t}^{(iii)}(Z_t | V_t)} + \eta \\ R_2 &\leq \frac{1}{n} \sum_{t=1}^n \log \frac{p_{Z_t | V_t}(Z_t | V_t)}{Q_{Z_t}^{(iv)}(Z_t)} + \eta \\ R_1 + R_2 &\leq \frac{1}{n} \sum_{t=1}^n \log \frac{W_1(Y_t | X_t)}{Q_{Y_t}^{(v)}(Y_t)} + \eta \Big\} + 5e^{-n\eta}, \quad (13) \end{aligned}$$

where for each  $t = 1, 2, \dots, n$ , the following probability and conditional probability distributions:

$$\left\{ Q_{Y_t | X_t Z_t U_t}^{(i)}, Q_{Z_t | X_t Y_t U_t}^{(ii)}, Q_{Y_t | U_t}^{(iii)}, Q_{Z_t | U_t}^{(iii)}, Q_{Z_t | V_t}^{(iii)}, Q_{Z_t}^{(iv)}, Q_{Y_t}^{(v)} \right\} \quad (14)$$

appearing in the first term in the right members of (13) have a property that we can choose their values arbitrary.

*Proof:* In (8), we choose  $Q_{Y^n | X^n Z^n L_n}^{(i)}$  so that

$$\begin{aligned} &Q_{Y^n | X^n Z^n L_n}^{(i)}(Y^n | X^n, Z^n, L_n) \\ &= \prod_{t=1}^n Q_{Y_t | X_t Y^{t-1} Z_{t+1}^n L_n}^{(i)}(Y_t | X_t, Y^{t-1}, Z_{t+1}^n, L_n) \\ &= \prod_{t=1}^n Q_{Y_t | X_t U_t}^{(i)}(Y_t | X_t, U_t). \end{aligned} \quad (15)$$

In (9), we choose  $Q_{Z^n | X^n Y^n L_n}^{(ii)}$  so that

$$\begin{aligned} &Q_{Z^n | X^n Y^n L_n}^{(ii)}(Z^n | X^n, Y^n, L_n) \\ &= \prod_{t=1}^n Q_{Z_t | X_t Y^{t-1} Z_{t+1}^n L_n}^{(ii)}(Z_t | X_t, Y^{t-1}, Z_{t+1}^n, L_n) \\ &= \prod_{t=1}^n Q_{Z_t | X_t U_t}^{(ii)}(Z_t | X_t, U_t). \end{aligned} \quad (16)$$



In (10), we have the following chain of equalities:

$$\begin{aligned}
& \frac{W_1^n(Y^n|X^n)}{Q_{Y^n|L_n}^{(iii)}(Y^n|L_n)} \\
&= \frac{W_1^n(Y^n|X^n)Q_{Z^n|L_n}^{(iii)}(Z^n|L_n)}{Q_{Y^n|L_n}^{(iii)}(Y^n|L_n)Q_{Z^n|L_n}^{(iii)}(Z^n|L_n)} \\
&= \prod_{t=1}^n \frac{Q_{Z_t|Y^{t-1}Z_{t+1}^nL_n}^{(iii)}(Z_t|Y^{t-1}, Z_{t+1}^n, L_n)}{Q_{Y_t|Y^{t-1}Z_{t+1}^nL_n}^{(iii)}(Z_t|Y^{t-1}, Z_{t+1}^n, L_n)} \\
&\quad \times \prod_{t=1}^n \frac{W_1(Y_t|X_t)}{Q_{Z_t|Z_{t+1}^nL_n}^{(iii)}(Z_t|Z_{t+1}^n, L_n)} \\
&= \prod_{t=1}^n \frac{W_1(Y_t|X_t)Q_{Z_t|U_t}^{(iii)}(Z_t|U_t)}{Q_{Y_t|U_t}^{(iii)}(Y_t|U_t)Q_{Z_t|V_t}^{(iii)}(Z_t|V_t)}. \tag{17}
\end{aligned}$$

In (11), we choose  $Q_{Z^n}^{(iv)}$  so that

$$Q_{Z^n}^{(iv)}(Z^n) = \prod_{t=1}^n Q_{Z_t}^{(iv)}(Z_t). \tag{18}$$

In (12), we choose  $Q_{Y^n}^{(v)}$  so that

$$Q_{Y^n}^{(v)}(Y^n) = \prod_{t=1}^n Q_{Y_t}^{(v)}(Y_t). \tag{19}$$

From Lemma 1 and (15)-(19), we have the bound (13) in Lemma 2. ■

For each  $t = 1, 2, \dots, n$ , let  $\underline{Q}_t$  be a set of all

$$\begin{aligned}
\underline{Q}_t = & (Q_{Y_t|X_tZ_tU_t}^{(i)}, Q_{Z_t|X_tY_tU_t}^{(ii)}, Q_{Y_t|U_t}^{(iii)}, Q_{Z_t|U_t}^{(iii)}, \\
& Q_{Z_t|V_t}^{(iii)}, Q_{Z_t}^{(iv)}, Q_{Y_t}^{(v)}).
\end{aligned}$$

Set

$$\underline{Q}^n \triangleq \prod_{t=1}^n \underline{Q}_t, \underline{Q}^n \triangleq \{\underline{Q}_t\}_{t=1}^n \in \underline{Q}^n.$$

To evaluate an upper bound of (13) in Lemma 2. We use the following lemma, which is well known as the Cramér's bound in the large deviation principle.

*Lemma 3:* For any real valued random variable  $A$  and any  $\theta > 0$ , we have

$$\Pr\{A \geq a\} \leq \exp[-(\lambda a - \log E[\exp(\theta A)])].$$

Here we define a quantity which serves as an exponential upper bound of  $P_c^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)})$ . Let  $\mathcal{P}^{(n)}(W_1, W_2)$  be a set of all probability distributions  $p_{L_n X^n Y^n Z^n}$  on  $\mathcal{L}_n \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$  having the form:

$$\begin{aligned}
& p_{L_n X^n Y^n Z^n}(l, x^n, y^n, z^n) \\
&= p_{L^n}(l) \prod_{t=1}^n p_{X_t|L_n X^{t-1}}(x_t|l, x^{t-1}) W_1(y_t|x_t) W_2(z_t|x_t).
\end{aligned}$$

For simplicity of notation we use the notation  $p^{(n)}$  for  $p_{L_n X^n Y^n Z^n} \in \mathcal{P}^{(n)}(W_1, W_2)$ . We assume that  $p_{U_t X_t Y_t Z_t} = p_{L_n X_t Y_t Z_t}$  is a marginal distribution induced by  $p^{(n)}$ . For  $t = 1, 2, \dots, n$ , we simply write  $p_t = p_{U_t X_t Y_t Z_t}$ . For each

$t = 1, 2, \dots, n$ , let  $\text{Proj}(\mathcal{U}_t \rightarrow \mathcal{V}_t)$  be a set of all projection  $\kappa_t$  from  $\mathcal{U}_t$  onto  $\mathcal{V}_t$ . For  $p^{(n)} \in \mathcal{P}^{(n)}(W_1, W_2)$ ,

$$\kappa^n = \{\kappa_t\}_{t=1}^n \in \prod_{t=1}^n \text{Proj}(\mathcal{U}_t \rightarrow \mathcal{V}_t),$$

and  $\underline{Q}^n \in \underline{Q}^n$ , we define

$$\begin{aligned}
& \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n) \\
& \triangleq \log E_{p^{(n)}} \left[ \left( \prod_{t=1}^n \left\{ \frac{W_1(Y_t|X_t)}{Q_{Y_t|X_tZ_tU_t}^{(i)}(Y_t|X_t, Z_t, U_t)} \right\}^{\alpha\theta} \right) \right. \\
& \quad \times \left( \prod_{t=1}^n \left\{ \frac{W_2(Z_t|X_t)}{Q_{Z_t|X_tY_tU_t}^{(ii)}(Z_t|X_t, Y_t, U_t)} \right\}^{\beta\theta} \right) \\
& \quad \times \left( \prod_{t=1}^n \left\{ \frac{W_1(Y_t|X_t)}{Q_{Y_t|U_t}^{(iii)}(Y_t|U_t)} \right\}^{\gamma\mu\theta} \right) \\
& \quad \times \left( \prod_{t=1}^n \left\{ \frac{Q_{Z_t|U_t}^{(iii)}(Z_t|U_t)}{Q_{Z_t|V_t}^{(iii)}(Z_t|V_t)} \right\}^{\gamma\mu\theta} \right) \\
& \quad \times \left( \prod_{t=1}^n \left\{ \frac{p_{Z_t|V_t}(Z_t|V_t)}{Q_{Z_t}^{(iv)}(Z_t)} \right\}^{\gamma\bar{\mu}\theta} \right) \\
& \quad \times \left. \left( \prod_{t=1}^n \left\{ \frac{W_1(Y_t|X_t)}{Q_{Y_t}^{(v)}(Y_t)} \right\}^{\bar{\gamma}\theta} \right) \right],
\end{aligned}$$

where for each  $t = 1, 2, \dots, n$ , the following probability and conditional probability distributions:

$$\left\{ Q_{Y_t|X_tZ_tU_t}^{(i)}, Q_{Z_t|X_tY_tU_t}^{(ii)}, Q_{Y_t|U_t}^{(iii)}, Q_{Z_t|U_t}^{(iii)}, Q_{Z_t|V_t}^{(iii)}, Q_{Z_t}^{(iv)}, Q_{Y_t}^{(v)} \right\} \tag{20}$$

appearing in the definition of  $\Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n)$  can be chosen arbitrary.

By Lemmas 2 and 3, we have the following proposition.

*Proposition 1:* For any  $\alpha, \beta, \gamma, \mu, \theta > 0$ , any  $\underline{Q}^n \in \underline{Q}^n$ , and any  $(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)})$  satisfying

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R_1, \frac{1}{n} \log |\mathcal{L}_n| \geq R_2,$$

we have

$$\begin{aligned}
& P_c^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}) \\
& \leq 6 \exp \left( -n[1 + \theta(1 + \alpha + \beta)]^{-1} \right. \\
& \quad \times \left\{ \theta[(\bar{\gamma} + \gamma\mu)R_1 + (\bar{\gamma} + \gamma\bar{\mu})R_2] \right. \\
& \quad \left. \left. - \frac{1}{n} \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n) \right\} \right).
\end{aligned}$$

*Proof:* We define six random variables  $A_i, i = 1, 2, \dots, 5$  by

$$\begin{aligned} A_1 &= \frac{1}{n} \sum_{t=1}^n \log \frac{W_1(Y_t|X_t)}{Q_{Y_t|X_t Z_t U_t}^{(i)}(Y_t|X_t Z_t U_t)}, \\ A_2 &= \frac{1}{n} \sum_{t=1}^n \log \frac{W_2(Z_t|X_t)}{Q_{Z_t|X_t Y_t U_t}^{(ii)}(Z_t|X_t Y_t U_t)}, \\ A_3 &= \frac{1}{n} \sum_{t=1}^n \log \frac{W_1(Y_t|X_t) Q_{Z_t|U_t}^{(iii)}(Y_t|U_t)}{Q_{Y_t|U_t}^{(iii)}(Y_t|U_t) Q_{Z_t|V_t}^{(iii)}(Z_t|V_t)} - R_1, \\ A_4 &= \frac{1}{n} \sum_{t=1}^n \log \frac{p_{Z_t|V_t}(Z_t|V_t)}{Q_{Z_t}^{(v)}(Z_t)} - R_2, \\ A_5 &= \frac{1}{n} \sum_{t=1}^n \log \frac{W_1(Y_t|X_t)}{Q_{Y_t}^{(v)}(Y_t)} - (R_1 + R_2). \end{aligned}$$

Then by Lemma 2, for any  $(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)})$  satisfying

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R_1, \frac{1}{n} \log |\mathcal{L}_n| \geq R_2,$$

we have

$$\begin{aligned} &P_c^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}) \\ &\leq p_{L_n X^n Y^n Z^n} \{A_i \geq -\eta \text{ for } i = 1, 2, 3, 4, 5\} \\ &\quad + 5e^{-n\eta} \\ &\leq p_{L_n X^n Y^n Z^n} \{\alpha A_1 + \beta A_2 + \gamma[\mu A_3 + \bar{\mu} A_4] \\ &\quad + \bar{\gamma} A_5 \geq -\eta[1 + \alpha + \beta]\} + 5e^{-n\eta} \\ &= p_{L_n X^n Y^n Z^n} \{A \geq a\} + 5e^{-n\eta}, \end{aligned} \quad (21)$$

where we set

$$\begin{aligned} A &\triangleq \alpha A_1 + \beta A_2 + \gamma[\mu A_3 + \bar{\mu} A_4] + \bar{\gamma} A_5, \\ a &\triangleq -\eta[1 + \alpha + \beta]. \end{aligned}$$

Applying Lemma 3 to the first term in the right member of (21), we have

$$\begin{aligned} &P_c^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}) \\ &\leq \exp \left[ -(\theta a - \log E_{p^{(n)}}[\exp(\theta A)]) \right] + 5e^{-n\eta} \\ &= \exp \left[ n \left\{ \theta(1 + \alpha + \beta)\eta \right. \right. \\ &\quad \left. \left. - \theta[(\bar{\gamma} + \gamma\mu)R_1 + (\bar{\gamma} + \gamma\bar{\mu})R_2] \right. \right. \\ &\quad \left. \left. + \frac{1}{n} \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n) \right\} \right] + 5e^{-n\eta}. \end{aligned} \quad (22)$$

We choose  $\eta$  so that

$$\begin{aligned} -\eta &= \theta(1 + \alpha + \beta)\eta \\ &\quad - \theta[(\bar{\gamma} + \gamma\mu)R_1 + (\bar{\gamma} + \gamma\bar{\mu})R_2] \\ &\quad + \frac{1}{n} \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n). \end{aligned} \quad (23)$$

Solving (23) with respect to  $\eta$ , we have

$$\begin{aligned} \eta &= [1 + \theta(1 + \alpha + \beta)]^{-1} \\ &\quad \times \left\{ \theta[(\bar{\gamma} + \gamma\mu)R_1 + (\bar{\gamma} + \gamma\bar{\mu})R_2] \right. \\ &\quad \left. - \frac{1}{n} \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n) \right\}. \end{aligned}$$

For this choice of  $\eta$  and (22), we have

$$\begin{aligned} &P_c^{(n)}(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)}) \leq 6e^{-n\eta} \\ &= 6 \exp \left[ -n \{1 + \theta(1 + \alpha + \beta)\}^{-1} \right. \\ &\quad \times \left\{ \theta[(\bar{\gamma} + \gamma\mu)R_1 + (\bar{\gamma} + \gamma\bar{\mu})R_2] \right. \\ &\quad \left. \left. - \frac{1}{n} \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n) \right\} \right], \end{aligned}$$

completing the proof. ■

Set

$$\begin{aligned} &\bar{\Omega}^{(\alpha, \beta, \gamma, \mu, \theta)}(W_1, W_2) \\ &\triangleq \sup_{n \geq 1} \max_{\substack{p^{(n)} \in \mathcal{P}^{(n)}(W_1, W_2), \\ \kappa^n \in \prod_{t=1}^n \text{Proj}(\mathcal{U}_t \rightarrow \mathcal{V}_t)}} \min_{\underline{Q}^n \in \underline{\mathcal{Q}}^n} 1 \\ &\quad \times \frac{1}{n} \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n). \end{aligned}$$

Then we have the following corollary from Proposition 1.

*Corollary 2:* For any positive  $R_1, R_2$  and for any positive  $\alpha, \beta, \gamma, \mu$ , and  $\theta$ , we have

$$\begin{aligned} &G(R_1, R_2|W_1, W_2) \\ &\geq \frac{\theta[(\bar{\gamma} + \gamma\mu)R_1 + (\bar{\gamma} + \gamma\bar{\mu})R_2] - \bar{\Omega}^{(\alpha, \beta, \gamma, \mu, \theta)}(W_1, W_2)}{1 + \theta(1 + \alpha + \beta)}. \end{aligned}$$

*Proof:* By the definition of  $\bar{\Omega}^{(\alpha, \beta, \gamma, \mu, \theta)}(W_1, W_2)$ , the definition of  $G^{(n)}(R_1, R_2|W_1, W_2)$ , and Proposition 1, we have

$$\begin{aligned} &G^{(n)}(R_1, R_2|W_1, W_2) \\ &\geq \frac{\theta[(\bar{\gamma} + \gamma\mu)R_1 + (\bar{\gamma} + \gamma\bar{\mu})R_2] - \bar{\Omega}^{(\alpha, \beta, \gamma, \mu, \theta)}(W_1, W_2)}{1 + \theta(1 + \alpha + \beta)} \\ &\quad - \frac{1}{n} \log 6, \end{aligned}$$

from which we have Corollary 2. ■

We shall call  $\bar{\Omega}^{(\alpha, \beta, \gamma, \mu, \theta)}(W_1, W_2)$  the communication potential. The above corollary implies that the analysis of  $\bar{\Omega}^{(\alpha, \beta, \gamma, \mu, \theta)}(W_1, W_2)$  leads to an establishment of a strong converse theorem for the BC.

In the following argument we drive an explicit upper bound of  $\bar{\Omega}^{(\alpha, \beta, \gamma, \mu, \theta)}(W_1, W_2)$ . To this end we use a new novel technique called *the recursive method*. This method is a powerful tool to drive a single letterized exponent function for rates below the rate distortion function. This method is also applicable to prove the exponential strong converse theorems for other network information theory problems [7], [8], [9].

For each  $t = 1, 2, \dots, n$ , define a function of  $(u_t, x_t, y_t, z_t) \in \mathcal{U}_t \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$  by

$$\begin{aligned} & f_{(p_{Z_t|V_t}, \kappa_t), \underline{Q}_t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t, z_t | u_t) \\ \triangleq & \left\{ \frac{W_1(y_t | x_t)}{Q_{Y_t|X_t Z_t U_t}^{(i)}(y_t | x_t, z_t, u_t)} \right\}^{\alpha\theta} \\ & \times \left\{ \frac{W_2(z_t | x_t)}{Q_{Z_t|X_t Y_t U_t}^{(ii)}(z_t | x_t, y_t, u_t)} \right\}^{\beta\theta} \\ & \times \left\{ \frac{W_1(y_t | x_t)}{Q_{Y_t|U_t}^{(iii)}(y_t | u_t)} \right\}^{\gamma\mu\theta} \\ & \times \left\{ \frac{Q_{Z_t|U_t}^{(iii)}(z_t | u_t)}{Q_{Z_t|V_t}^{(iii)}(z_t | v_t)} \right\}^{\gamma\mu\theta} \left\{ \frac{p_{Z_t|V_t}(z_t | v_t)}{Q_{Z_t}^{(iv)}(z_t)} \right\}^{\gamma\bar{\mu}\theta} \\ & \times \left\{ \frac{W_1(y_t | x_t)}{Q_{Y_t}^{(v)}(y_t)} \right\}^{\bar{\gamma}\theta}. \end{aligned}$$

Set

$$\mathcal{F}_t \triangleq (p_{Z_t|V_t}, \kappa_t, \underline{Q}_t), \quad \mathcal{F}^t \triangleq \{\mathcal{F}_i\}_{i=1}^t.$$

By definition we have

$$\begin{aligned} & \exp \left\{ \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n) \right\} \\ = & \sum_{l, z^n, x^n, y^n} p_{L_n Z^n}(l, z^n) p_{X^n Y^n | L_n Z^n}(x^n, y^n | l, z^n) \\ & \times \prod_{t=1}^n f_{\mathcal{F}_t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t, z_t | u_t). \end{aligned} \quad (24)$$

For each  $t = 1, 2, \dots, n$ , we define a conditional probability distribution of  $(X^t, Y^t)$  given  $(L_n, Z^n)$  by

$$\begin{aligned} & p_{X^t Y^t | L_n Z^n, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)} \\ \triangleq & \left\{ p_{X^t Y^t | L_n Z^n, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x^t, y^t | l, z^n) \right\}_{(x^t, y^t, l, z^n) \in \mathcal{X}^t \times \mathcal{Y}^t \times \mathcal{L}_n \times \mathcal{Z}^n}, \\ & p_{X^t Y^t | L_n Z^n, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x^t, y^t | l, z^n) \\ \triangleq & C_t^{-1}(l, z^n) p_{X^t Y^t | L_n Z^n}(x^t, y^t | l, z^n) \\ & \times \prod_{i=1}^t f_{\mathcal{F}_i}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_i, y_i, z_i | u_i), \end{aligned}$$

where

$$\begin{aligned} C_t(l, z^n) \triangleq & \sum_{x^t, y^t} p_{X^t Y^t | L_n Z^n}(x^t, y^t | l, z^n) \\ & \times \prod_{i=1}^t f_{\mathcal{F}_i}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_i, y_i, z_i | u_i) \end{aligned} \quad (25)$$

are constants for normalization. For  $t = 1, 2, \dots, n$ , define

$$\Phi_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) \triangleq C_t(l, z^n) C_{t-1}^{-1}(l, z^n), \quad (26)$$

where we define  $C_0(l, z^n) = 1$  for  $(l, z^n) \in \mathcal{L}_n \times \mathcal{Z}^n$ . Then we have the following lemma.

**Lemma 4:** For each  $t = 1, 2, \dots, n$ , and for any  $(l, z^n, x^t, y^t) \in \mathcal{L}_n \times \mathcal{Z}^n \times \mathcal{X}^t \times \mathcal{Y}^t$ , we have

$$\begin{aligned} & p_{X^t Y^t | L_n Z^n, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x^t, y^t | l, z^n) = (\Phi_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n))^{-1} \\ & \times p_{X^{t-1} Y^{t-1} | L_n Z^n, \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(x^{t-1}, y^{t-1} | l, z^n) \\ & \times p_{X_t Y_t | L_n X^{t-1} Z^n}(x_t, y_t | l, x^{t-1}, y^{t-1}, z^n) \\ & \times f_{\mathcal{F}_t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t, z_t | u_t), \end{aligned} \quad (27)$$

$$\begin{aligned} & \Phi_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) \\ = & \sum_{x^t, y^t} p_{X^{t-1} Y^{t-1} | L_n Z^n, \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(x^{t-1}, y^{t-1} | l, z^n) \\ & \times p_{X_t Y_t | L_n X^{t-1} Z^n}(x_t, y_t | l, x^{t-1}, y^{t-1}, z^n) \\ & \times f_{\mathcal{F}_t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t, z_t | u_t). \end{aligned} \quad (28)$$

Furthermore, we have

$$\begin{aligned} & \exp \left\{ \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n) \right\} \\ = & \sum_{l, z^n} p_{L_n Z^n}(l, z^n) \prod_{t=1}^n \Phi_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n). \end{aligned} \quad (29)$$

The equality (29) in Lemma 4 is obvious from (24), (25), and (26). Proofs of (27) and (28) in this lemma are given in Appendix E. Next we define the probability distribution

$$p_{L_n Z^n, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)} = \left\{ p_{L_n Z^n, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) \right\}_{(l, z^n) \in \mathcal{L}_n \times \mathcal{Z}^n}$$

of the random variable  $(L_n, Z^n)$  taking values in  $\mathcal{L}_n \times \mathcal{Z}^n$  by

$$\begin{aligned} & p_{L_n Z^n, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) \\ = & \tilde{C}_t^{-1} p_{L_n Z^n}(l, z^n) \prod_{i=1}^t \Phi_{i, \mathcal{F}^i}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n), \end{aligned} \quad (30)$$

where  $\tilde{C}_t$  is a constant for normalization given by

$$\tilde{C}_t = \sum_{l, z^n} p_{L_n Z^n}(l, z^n) \prod_{i=1}^t \Phi_{i, \mathcal{F}^i}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n).$$

For  $t = 1, 2, \dots, n$ , define

$$\Lambda_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)} \triangleq \tilde{C}_t \tilde{C}_{t-1}^{-1}, \quad (31)$$

where we define  $\tilde{C}_0 = 1$ . Then we have the following.

**Lemma 5:**

$$\Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n) = \sum_{t=1}^n \log \Lambda_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}, \quad (32)$$

$$\begin{aligned} & \Lambda_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)} \\ = & \sum_{l, z^n} p_{L_n Z^n, \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) \Phi_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) \\ = & \sum_{l, z^n} \sum_{x^t, y^t} p_{L_n Z^n, \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) \\ & \times p_{X^{t-1} Y^{t-1} | L_n Z^n, \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(x^{t-1}, y^{t-1} | l, z^n) \\ & \times p_{X_t Y_t | X^{t-1} Y^{t-1} L_n Z^n}(x_t, y_t | x^{t-1}, y^{t-1}, l, z^n) \\ & \times f_{\mathcal{F}_t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t, z_t | u_t). \end{aligned} \quad (33)$$



*Proof:* By the equality (29) in Lemma 4, we have

$$\begin{aligned} & \exp \left\{ \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n) \right\} \\ &= \tilde{C}_n = \prod_{t=1}^n \tilde{C}_t \tilde{C}_{t-1}^{-1} \stackrel{(a)}{=} \prod_{t=1}^n \Lambda_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}. \end{aligned} \quad (34)$$

Step (a) follows from the definition (31) of  $\Lambda_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}$ . From (34), we have (32) in Lemma 5. We next prove (33) in Lemma 5. Multiplying  $\Lambda_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)} = \tilde{C}_t / \tilde{C}_{t-1}$  to both sides of (30), we have

$$\Lambda_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)} p_{L_n Z^n; \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) \quad (35)$$

$$\begin{aligned} &= \tilde{C}_{t-1}^{-1} p_{L_n Z^n}(l, z^n) \prod_{i=1}^t \Phi_{i, \mathcal{F}^i}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n), \\ &= p_{L_n Z^n; \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) \Phi_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n). \end{aligned} \quad (36)$$

Taking summations of (35) and (36) with respect to  $(l, z^n)$ , we have (33) in Lemma 5. ■

The following proposition is a mathematical core to prove our main result.

*Proposition 2:* For  $\theta \in (0, [(2 - 3\mu)\gamma]^{-1})$ , set

$$\lambda = \frac{\theta}{1 - (2 - 3\mu)\gamma\theta} \Leftrightarrow \theta = \frac{\lambda}{1 + (2 - 3\mu)\gamma\lambda}. \quad (37)$$

Then, for any positive  $\alpha, \beta, \gamma, \mu$ , and any  $\theta \in (0, [(2 - 3\mu)\gamma]^{-1})$ , we have

$$\bar{\Omega}^{(\alpha, \beta, \gamma, \mu, \theta)}(W_1, W_2) \leq \frac{\Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)}{1 + (2 - 3\mu)\gamma\lambda}.$$

*Proof:* Set

$$\begin{aligned} \hat{\mathcal{Q}}_n &\triangleq \{q = q_{UXYZ} : |\mathcal{U}| \leq |\mathcal{L}_n| |\mathcal{Y}^{n-1}| |\mathcal{Z}^{n-1}|\}, \\ \hat{\Omega}_n^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2) &\triangleq \min_{q \in \hat{\mathcal{Q}}_n} \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2). \end{aligned}$$

Set

$$\begin{aligned} & p_{L_n X_t Y^t Z_t; \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, x_t, y^t, z_t^n) \\ &= p_{U_t X_t Y_t Z_t; \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(u_t, x_t, y_t, z_t) \\ &\triangleq \sum_{x^{t-1}, z^{t-1}} p_{L_n Z^n; \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) \\ &\quad \times p_{X^{t-1} Y^{t-1} | L_n Z^n; \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(x^{t-1}, y^{t-1} | l, z^n) \\ &\quad \times p_{X_t Y_t | X^{t-1} Y^{t-1} L_n Z^n}(x_t, y_t | x^{t-1}, y^{t-1}, l, z^n). \end{aligned} \quad (38)$$

Then by Lemma 5, we have

$$\begin{aligned} \Lambda_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)} &= \sum_{u_t, x_t, y_t, z_t} p_{U_t X_t Y_t Z_t; \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(u_t, x_t, y_t, z_t) \\ &\quad \times f_{\mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t, z_t | u_t). \end{aligned}$$

For each  $t = 1, 2, \dots, n$ , we choose  $q_t = q_{U_t X_t Y_t Z_t}$  so that

$$q_{U_t X_t Y_t Z_t}(u_t, x_t, y_t, z_t) = p_{U_t X_t Y_t Z_t; \mathcal{F}^{t-1}}^{(\alpha, \beta, \gamma, \mu, \theta)}(u_t, x_t, y_t, z_t)$$

and choose the following probability and conditional probability distributions:

$$\left\{ \begin{aligned} & Q_{Y_t | X_t Z_t U_t}^{(i)}, Q_{Z_t | X_t Y_t U_t}^{(ii)}, Q_{Y_t | U_t}^{(iii)}, Q_{Z_t | U_t}^{(iii)} \\ & Q_{Z_t}^{(iv)}, Q_{Y_t}^{(v)} \end{aligned} \right\}$$

appearing in

$$\begin{aligned} & f_{\mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t, z_t | u_t) \\ &= \left\{ \frac{W_1(y_t | x_t)}{Q_{Y_t | X_t Z_t U_t}^{(i)}(y_t | x_t, z_t, u_t)} \right\}^{\alpha\theta} \\ &\quad \times \left\{ \frac{W_2(z_t | x_t)}{Q_{Z_t | X_t Y_t U_t}^{(ii)}(z_t | x_t, y_t, u_t)} \right\}^{\beta\theta} \left\{ \frac{W_1(y_t | x_t)}{Q_{Y_t | U_t}^{(iii)}(y_t | u_t)} \right\}^{\gamma\mu\theta} \\ &\quad \times \left\{ \frac{Q_{Z_t | U_t}^{(iii)}(z_t | u_t)}{Q_{Z_t | V_t}^{(iii)}(z_t | v_t)} \right\}^{\gamma\mu\theta} \left\{ \frac{p_{Z_t | V_t}(z_t | v_t)}{Q_{Z_t}^{(iv)}(z_t)} \right\}^{\gamma\bar{\mu}\theta} \\ &\quad \times \left\{ \frac{W_1(y_t | x_t)}{Q_{Y_t}^{(v)}(y_t)} \right\}^{\bar{\gamma}\theta} \end{aligned}$$

such that they are the distributions induced by  $q_{U_t X_t Y_t Z_t}$ . Then for each  $t = 1, 2, \dots, n$ , we have the following chain of inequalities:

$$\begin{aligned} & \Lambda_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)} \\ &= E_{q_t} \left[ \left\{ \frac{W_1^{\alpha\theta}(Y_t | X_t)}{q_{Y_t | X_t Z_t U_t}^{\alpha\theta}(Y_t | X_t, Z_t, U_t)} \right. \right. \\ &\quad \times \frac{W_2^{\beta\theta}(Z_t | X_t)}{q_{Z_t | X_t Y_t U_t}^{\beta\theta}(Z_t | X_t, Y_t, U_t)} \frac{W_1^{\gamma\mu\theta}(Y_t | X_t)}{q_{Y_t | U_t}^{\gamma\mu\theta}(Y_t | U_t)} \\ &\quad \times \frac{q_{Z_t | U_t}^{\gamma\bar{\mu}\theta}(Z_t | U_t)}{q_{Z_t}^{\gamma\bar{\mu}\theta}(Z_t)} \frac{W_1^{\bar{\gamma}\theta}(Y_t | X_t)}{q_{Y_t}^{\bar{\gamma}\theta}(Y_t)} \left. \right\} \\ &\quad \times \left\{ \frac{p_{Z_t | V_t}(Z_t | V_t)}{q_{Z_t | V_t}^{\gamma\bar{\mu}\theta}(Z_t | V_t)} \right\} \left\{ \frac{q_{Z_t | V_t}^{\gamma(\bar{\mu}-\mu)\theta}(Z_t | V_t)}{q_{Z_t | U_t}^{\gamma(\bar{\mu}-\mu)\theta}(Z_t | U_t)} \right\} \left. \right] \\ &\stackrel{(a)}{\leq} \left( E_{q_t} \left[ \left\{ \frac{W_1^{\alpha\theta}(Y_t | X_t)}{q_{Y_t | X_t Z_t U_t}^{\alpha\theta}(Y_t | X_t, Z_t, U_t)} \right. \right. \right. \\ &\quad \times \frac{W_2^{\beta\theta}(Z_t | X_t)}{q_{Z_t | X_t Y_t U_t}^{\beta\theta}(Z_t | X_t, Y_t, U_t)} \frac{W_1^{\gamma\mu\theta}(Y_t | X_t)}{q_{Y_t | U_t}^{\gamma\mu\theta}(Y_t | U_t)} \\ &\quad \times \frac{q_{Z_t | U_t}^{\gamma\bar{\mu}\theta}(Z_t | U_t)}{q_{Z_t}^{\gamma\bar{\mu}\theta}(Z_t)} \frac{W_1^{\bar{\gamma}\theta}(Y_t | X_t)}{q_{Y_t}^{\bar{\gamma}\theta}(Y_t)} \left. \right\}^{1 - (2 - 3\mu)\gamma\theta} \left. \right]^{1 - (2 - 3\mu)\gamma\theta} \\ &\quad \times \left\{ E_{q_t} \left[ \frac{p_{Z_t | V_t}(Z_t | V_t)}{q_{Z_t | V_t}^{\gamma\bar{\mu}\theta}(Z_t | V_t)} \right] \right\}^{\gamma\bar{\mu}\theta} \\ &\quad \times \left\{ E_{q_t} \left[ \frac{q_{Z_t | V_t}^{\gamma(\bar{\mu}-\mu)\theta}(Z_t | V_t)}{q_{Z_t | U_t}^{\gamma(\bar{\mu}-\mu)\theta}(Z_t | U_t)} \right] \right\}^{\gamma(\bar{\mu}-\mu)\theta} \\ &= \exp \left\{ [1 - (2 - 3\mu)\gamma\theta] \right. \\ &\quad \times \Omega^{(\alpha, \beta, \gamma, \mu, \frac{\theta}{1 - (2 - 3\mu)\gamma\theta})}(q_t | W_1, W_2) \left. \right\} \end{aligned}$$

$$\begin{aligned}
&\stackrel{(b)}{=} \exp \left\{ \frac{\Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(q_t | W_1, W_2)}{1 + (2 - 3\mu)\gamma\lambda} \right\} \\
&\stackrel{(c)}{\leq} \exp \left\{ \frac{\hat{\Omega}_n^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)}{1 + (2 - 3\mu)\gamma\lambda} \right\} \\
&\stackrel{(d)}{=} \exp \left\{ \frac{\Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)}{1 + (2 - 3\mu)\gamma\lambda} \right\}. \tag{39}
\end{aligned}$$

Step (a) follows from Hölder's inequality. Step (b) follows from (37). Step (c) follows from  $q_t \in \hat{\mathcal{P}}_n(W_1, W_2)$  and the definition of  $\hat{\Omega}_n^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)$ . Step (d) follows from Lemma 6 in Appendix A. To prove this lemma we bound the cardinality  $|\mathcal{U}|$  appearing in the definition of  $\hat{\Omega}_n^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)$  to show that the bound  $|\mathcal{U}| \leq |\mathcal{Y}| + |\mathcal{Z}| - 1$  is sufficient to describe  $\hat{\Omega}_n^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)$ . Hence we have the following:

$$\begin{aligned}
&\min_{q^n \in \mathcal{Q}^n} \frac{1}{n} \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n) \\
&\leq \frac{1}{n} \Omega^{(\alpha, \beta, \gamma, \mu, \theta)}(p^{(n)}, \kappa^n, \underline{Q}^n) \stackrel{(a)}{=} \frac{1}{n} \sum_{t=1}^n \log \Lambda_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)} \\
&\stackrel{(b)}{\leq} \frac{\Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)}{1 + (2 - 3\mu)\gamma\lambda}. \tag{40}
\end{aligned}$$

Step (a) follows from (32) in Lemma 5. Step (b) follows from (39). Since (40) holds for any  $n \geq 1$  and any  $p^{(n)} \in \mathcal{P}^{(n)}(W_1, W_2)$ , we have

$$\bar{\Omega}^{(\alpha, \beta, \gamma, \mu, \theta)}(W_1, W_2) \leq \frac{\Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)}{1 + (2 - 3\mu)\gamma\lambda},$$

completing the proof.  $\blacksquare$

*Proof of Theorem 3:* For  $\theta \in (0, [(2 - 3\mu)\gamma]^{-1})$ , set

$$\lambda = \frac{\theta}{1 - (2 - 3\mu)\gamma\theta} \Leftrightarrow \theta = \frac{\lambda}{1 + (2 - 3\mu)\gamma\lambda}. \tag{41}$$

Then we have the following:

$$\begin{aligned}
&G(R_1, R_2 | W_1, W_2) \\
&\stackrel{(a)}{\geq} \frac{\theta[(\gamma\mu + \bar{\gamma})R_1 + (\gamma\bar{\mu} + \bar{\gamma})R_2] - \bar{\Omega}^{(\alpha, \beta, \gamma, \mu, \theta)}(W_1, W_2)}{1 + \theta[1 + \alpha + \beta]} \\
&\stackrel{(b)}{\geq} \frac{\lambda[(\gamma\mu + \bar{\gamma})R_1 + (\gamma\bar{\mu} + \bar{\gamma})R_2] - \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)}{1 + (2 - 3\mu)\gamma\lambda} \\
&\quad \frac{\lambda[1 + \alpha + \beta]}{1 + (2 - 3\mu)\gamma\lambda} \\
&= \frac{\lambda[(\gamma\mu + \bar{\gamma})R_1 + (\gamma\bar{\mu} + \bar{\gamma})R_2] - \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)}{1 + \lambda[1 + \alpha + \beta + (2 - 3\mu)\gamma]} \\
&= F^{(\alpha, \beta, \gamma, \mu, \lambda)}(R_1, R_2 | W_1, W_2). \tag{42}
\end{aligned}$$

Step (a) follows from Corollary 2. Step (b) follows from Proposition 2 and (41). Since (42) holds for any positive  $\alpha, \beta, \mu, \nu$ , and  $\lambda$ , we have

$$G(R_1, R_2 | W_1, W_2) \geq F(R_1, R_2 | W_1, W_2).$$

Thus (6) in Theorem 3 is proved. The inclusion  $\mathcal{R}(W_1, W_2) \subseteq \bar{\mathcal{R}}(W_1, W_2)$  is obvious from this bound.  $\blacksquare$

## APPENDIX

### A. Cardinality Bound on Auxiliary Random Variables

In this appendix we prove the cardinality bounds on auxiliary random variables appearing in this paper. We first prove Property 2 part a). Observe that

$$p_X(x) = \sum_{u \in \mathcal{U}} p_U(u) p_{X|U}(x|u), \tag{43}$$

$$\begin{aligned}
p_Y(y) &= \sum_{u \in \mathcal{U}} \sum_{x \in \mathcal{X}} p_U(u) W_1(y|x) p_{X|U}(x|u), \\
p_Z(z) &= \sum_{u \in \mathcal{U}} \sum_{x \in \mathcal{X}} p_U(u) W_2(z|x) p_{X|U}(x|u),
\end{aligned} \tag{44}$$

$$\begin{aligned}
&\gamma[\mu I_p(X; Y|U) + \bar{\mu} I_p(U; Z)] + \bar{\gamma} I_p(X; Y) \\
&= \sum_{u \in \mathcal{U}} p_U(u) [\pi_1(p_{X|U}(\cdot|u)) + \pi_2(p_{X|U}(\cdot|u))], \tag{45}
\end{aligned}$$

where we set

$$\begin{aligned}
\pi_1(p_{X|U}(\cdot|u)) &\triangleq \sum_{(x, y) \in \mathcal{X} \times \mathcal{Y}} p_{X|U}(x|u) W_1(y|x) \\
&\quad \times \log \left\{ \frac{\left[ \frac{W_1(y|x)}{\sum_{\tilde{x} \in \mathcal{X}} W_1(y|\tilde{x}) p_X(\tilde{x})} \right]^{\bar{\gamma} + \gamma\bar{\mu}}}{\left[ \frac{\sum_{\tilde{x} \in \mathcal{X}} W_1(y|\tilde{x}) p_{X|U}(\tilde{x}|u)}{p_Y(y)} \right]^{\gamma\bar{\mu}}} \right\}, \\
\pi_2(p_{X|U}(\cdot|u)) &\triangleq \sum_{(x, z) \in \mathcal{X} \times \mathcal{Z}} p_{X|U}(x|u) W_2(z|x) \\
&\quad \times \log \left\{ \left[ \frac{\sum_{\tilde{x} \in \mathcal{X}} W_2(z|\tilde{x}) p_{X|U}(\tilde{x}|u)}{p_Z(z)} \right]^{\gamma\bar{\mu}} \right\}.
\end{aligned}$$

*Proof of Property 2 part a):* We bound the cardinality  $|\mathcal{U}|$  of  $U$  to show that the bound  $|\mathcal{U}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}| + |\mathcal{Z}| - 1\}$  is sufficient to describe  $C^{(\mu)}(W_1, W_2)$ . We first derive a sufficient value of  $|\mathcal{U}|$  to express  $|\mathcal{X}| - 1$  values of (43) and (45). Note that by (44) the quantities  $p_Y(\cdot)$  and  $p_Z(\cdot)$  appearing in the above definitions of  $\pi_i(p_{X|U}(\cdot|u))$ ,  $i = 1, 2$ , are regarded as constants under (43). For each  $u \in \mathcal{U}$ ,  $\pi_i(p_{X|U}(\cdot|u))$ ,  $i = 1, 2$  is a continuous function of  $p_{X|U}(\cdot|u)$ . Then by the support lemma,

$$|\mathcal{U}| \leq |\mathcal{X}| - 1 + 1 = |\mathcal{X}|$$

is sufficient to express  $|\mathcal{X}| - 1$  values of (43) and one value of (45). We next derive a sufficient value of  $|\mathcal{U}|$  to express  $|\mathcal{Y}| + |\mathcal{Z}| - 2$  values of (44) and (45). Note that the quantities  $p_Y(\cdot)$  and  $p_Z(\cdot)$  appearing in the above definitions of  $\pi_i(p_{X|U}(\cdot|u))$ ,  $i = 1, 2$ , are regarded as constants under (44). For each  $u \in \mathcal{U}$ ,  $\pi_i(p_{X|U}(\cdot|u))$ ,  $i = 1, 2$  is a continuous function of  $p_{X|U}(\cdot|u)$ . Then by the support lemma,

$$|\mathcal{U}| \leq |\mathcal{Y}| + |\mathcal{Z}| - 2 + 1 = |\mathcal{Y}| + |\mathcal{Z}| - 1$$

is sufficient to express  $|\mathcal{Y}| + |\mathcal{Z}| - 2$  values of (44) and one value of (45). ■

Next we prove the following lemma.

**Lemma 6:** For each integer  $n \geq 2$ , we define

$$\begin{aligned} & \hat{\Omega}_n^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2) \\ & \triangleq \max_{\substack{q=q_{UXYZ}: \\ |\mathcal{U}| \leq |\mathcal{L}_n| |\mathcal{Y}|^{n-1} |\mathcal{Z}|^{n-1}}} \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2), \\ & \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2) \\ & \triangleq \max_{\substack{q=q_{UXYZ}: \\ |\mathcal{U}| \leq |\mathcal{Y}| + |\mathcal{Z}| - 1}} \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2). \end{aligned}$$

Then we have

$$\hat{\Omega}^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2) = \Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2).$$

*Proof:* We bound the cardinality  $|\mathcal{U}|$  of  $U$  to show that the bound  $|\mathcal{U}| \leq |\mathcal{Y}| + |\mathcal{Z}| - 1$  is sufficient to describe  $\hat{\Omega}_n^{(\alpha, \beta, \gamma, \mu, \lambda)}(W_1, W_2)$ . Observe that

$$\left. \begin{aligned} q_Y(y) &= \sum_{u \in \mathcal{U}} q_U(u) q_{Y|U}(y|u), \\ q_Z(z) &= \sum_{u \in \mathcal{U}} q_U(u) q_{Z|U}(z|u), \end{aligned} \right\} \quad (46)$$

$$\begin{aligned} & \Lambda^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2) \\ &= \sum_{u \in \mathcal{U}} q_U(u) \zeta^{(\alpha, \beta, \gamma, \mu, \lambda)}(q_{XYZ|U}(\cdot|u)), \end{aligned} \quad (47)$$

where we set

$$\begin{aligned} & \zeta^{(\alpha, \beta, \gamma, \mu, \lambda)}(q_{XYZ|U}(\cdot, \cdot, \cdot|u)) \\ & \triangleq \sum_{(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}} q_{XYZ|U}(x, y, z|u) \\ & \quad \times \exp \left\{ \lambda \omega_q^{(\alpha, \beta, \gamma, \mu)}(x, y, z|u) \right\}. \end{aligned}$$

For the quantities  $q_Y(\cdot)$  and  $q_Z(\cdot)$  contained in the forms of  $\zeta^{(\alpha, \beta, \gamma, \mu, \lambda)}(q_{XYZ|U}(\cdot|u))$ ,  $u \in \mathcal{U}$ , we regard them as constants under (46). For each  $u \in \mathcal{U}$ ,  $\zeta^{(\alpha, \beta, \gamma, \mu, \lambda)}(q_{XYZ|U}(\cdot|u))$  is a continuous function of  $q_{XYZ|U}(\cdot, \cdot, \cdot|u)$ . Then by the support lemma,

$$|\mathcal{U}| \leq |\mathcal{Y}| + |\mathcal{Z}| - 2 + 1 = |\mathcal{Y}| + |\mathcal{Z}| - 1$$

is sufficient to express  $|\mathcal{Y}| + |\mathcal{Z}| - 2$  values of (46) and one value of (47). ■

### B. Supporting Hyperplane Expression of the Capacity Region

In this appendix we prove Property 2 part b). From Property 1 part c), we have the following lemma.

**Lemma 7:** Suppose that  $(\hat{r}_1, \hat{r}_2, \hat{r}_3)$  does not belong to  $\mathcal{C}_{\text{ext}}(W_1, W_2)$ . Then there exist  $\epsilon, \gamma^*, \mu^* > 0$  such that for any  $(r_1, r_2, r_3) \in \mathcal{C}_{\text{ext}}(W_1, W_2)$  we have

$$\begin{aligned} & \gamma^* \mu^* (r_1 - \hat{r}_1) + \gamma^* \bar{\mu}^* (r_2 - \hat{r}_2) \\ & + \bar{\gamma}^* (r_3 - \hat{r}_3) + \epsilon \leq 0. \end{aligned}$$

Proof of this lemma is omitted here. This lemma will be used to prove Property 2 part b).

*Proof of Property 2 part b):* We first recall the following definitions of  $\mathcal{P}(W_1, W_2)$ ,  $\mathcal{P}_{\text{sh}}(W_1, W_2)$ , and  $\mathcal{Q}$ :

$$\begin{aligned} & \mathcal{P}(W_1, W_2) \\ & \triangleq \{p = p_{UXYZ} : |\mathcal{U}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}| + |\mathcal{Z}|\} + 1, \\ & \quad p_{Y|X} = W_1, p_{Z|X} = W_2, \\ & \quad Y \leftrightarrow X \leftrightarrow Z, U \leftrightarrow X \leftrightarrow (Y, Z)\}, \\ & \mathcal{P}_{\text{sh}}(W_1, W_2) \\ & \triangleq \{p = p_{UXYZ} : |\mathcal{U}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}| + |\mathcal{Z}| - 1\}, \\ & \quad p_{Y|X} = W_1, p_{Z|Y} = W_2, \\ & \quad Y \leftrightarrow X \leftrightarrow Z, U \leftrightarrow X \leftrightarrow (Y, Z)\}, \\ & \mathcal{Q} \triangleq \{q = q_{UXYZ} : |\mathcal{U}| \leq |\mathcal{Y}| + |\mathcal{Z}| - 1\}. \end{aligned}$$

We first show that  $\mu \in [0, 1/2]$  is sufficient to describe  $\mathcal{C}_{\text{ext}}(W_1, W_2)$ . The supporting hyperplane of  $\mathcal{C}_{\text{ext}}(W_1, W_2)$  with a normal vector  $(\gamma\mu, \gamma\bar{\mu}, \bar{\gamma})$  can be mapped to that of  $\mathcal{C}(W_1, W_2)$  with  $(\gamma\mu + \bar{\gamma}, \gamma\bar{\mu} + \bar{\gamma})$ . Hence by Property 1 part d), we have

$$\gamma\bar{\mu} + \bar{\gamma} - (\gamma\mu + \bar{\gamma}) = \gamma(1 - 2\mu) \geq 0.$$

We next prove  $\mathcal{C}_{\text{ext,sh}}(W_1, W_2) \subseteq \mathcal{C}_{\text{ext}}(W_1, W_2)$ . We assume that  $(\hat{r}_1, \hat{r}_2, \hat{r}_3) \notin \mathcal{C}_{\text{ext}}(W_1, W_2)$ . Then by Lemma 7, there exist  $\epsilon, \gamma^* \in [0, 1], \mu^* \in [0, 1/2]$  such that for any  $(r_1, r_2, r_3) \in \mathcal{C}_{\text{ext}}(W_1, W_2)$  we have

$$\begin{aligned} & \gamma^* \mu^* (r_1 - \hat{r}_1) + \gamma^* \bar{\mu}^* (r_2 - \hat{r}_2) \\ & + \bar{\gamma}^* (r_3 - \hat{r}_3) + \epsilon \leq 0. \end{aligned}$$

Then we have

$$\begin{aligned} & \gamma^* (\mu^* \hat{r}_1 + \bar{\mu}^* \hat{r}_2) + \bar{\gamma}^* \hat{r}_3 \\ & \geq \max_{\substack{(r_1, r_2, r_3) \\ \in \mathcal{C}(W_1, W_2)}} \{ \gamma^* (\mu^* r_1 + \bar{\mu}^* r_2) + \bar{\gamma}^* r_3 \} + \epsilon \\ & \stackrel{(a)}{=} \max_{p \in \mathcal{P}(W_1, W_2)} \{ \gamma^* \mu^* I_p(X; Y|U) + \gamma^* \bar{\mu}^* I_p(U; Z) \\ & \quad + \bar{\gamma}^* I_p(X; Y) \} + \epsilon \\ & = C^{(\gamma^*, \mu^*)}(W_1, W_2) + \epsilon. \end{aligned} \quad (48)$$

Step (a) follows from the definition of  $\mathcal{C}_{\text{ext}}(W_1, W_2)$ . The bound (48) implies that  $(\hat{r}_1, \hat{r}_2, \hat{r}_3) \notin \mathcal{C}_{\text{ext,sh}}(W_1, W_2)$ . Thus  $\mathcal{C}_{\text{ext,sh}}(W_1, W_2) \subseteq \mathcal{C}_{\text{ext}}(W_1, W_2)$  is proved. To complete the proof it suffices to prove the following inclusions:

$$\mathcal{C}_{\text{ext}}(W_1, W_2) \subseteq \tilde{\mathcal{C}}_{\text{ext,sh}}(W_1, W_2) \subseteq \mathcal{C}_{\text{ext,sh}}(W_1, W_2).$$

We first prove  $\mathcal{C}_{\text{ext}}(W_1, W_2) \subseteq \tilde{\mathcal{C}}_{\text{ext,sh}}(W_1, W_2)$ . We assume that  $(R_1, R_2) \in \mathcal{C}_{\text{ext}}(W_1, W_2)$ . Then there exists  $p \in \mathcal{P}(W_1, W_2)$  such that

$$\left. \begin{aligned} & r_1 \leq I_p(X; Y|U), r_2 \leq I_p(U; Z), \\ & r_3 \leq I_p(X; Y). \end{aligned} \right\} \quad (49)$$

Then, for  $(r_1, r_2, r_3) \in \mathcal{C}_{\text{ext}}(W_1, W_2)$ , we have the following chain of inequalities:

$$\begin{aligned}
& \gamma(\mu r_1 + \bar{\mu} r_2) + \bar{\gamma} r_3 \\
& \stackrel{(a)}{\leq} \gamma[\mu I_p(X; Y|U) + \bar{\mu} I_p(U; Z)] + \bar{\gamma} I_p(X; Y) \\
& \leq \max_{p \in \mathcal{P}(W_1, W_2)} \{ \gamma \mu I_p(X; Y|U) + \gamma \bar{\mu} I_p(U; Z) \\
& \quad + \bar{\gamma} I_p(X; Y) \} \\
& \stackrel{(b)}{=} \max_{p \in \mathcal{P}_{\text{sh}}(W_1, W_2)} \{ \gamma[\mu I_p(X; Y|U) + \bar{\mu} I_p(U; Z)] \\
& \quad + \bar{\gamma} I_p(X; Y) \} \\
& \stackrel{(c)}{=} \max_{p \in \mathcal{P}_{\text{sh}}(W_1, W_2)} \{ -\alpha D(p_{Y|XZU} \| W_1 | p_{XZU}) \\
& \quad - \beta D(p_{Z|XYU} \| W_2 | p_{XYU}) \\
& \quad + \gamma[\mu I_p(X; Y|U) + \bar{\mu} I_p(U; Z)] \\
& \quad + \bar{\gamma} I_p(X; Y) \} \\
& \leq \max_{p \in \mathcal{Q}} \{ -\alpha D(p_{Y|XZU} \| W_1 | p_{XZU}) \\
& \quad - \beta D(p_{Z|XYU} \| W_2 | p_{XYU}) \\
& \quad + \gamma[\mu I_p(X; Y|U) + \bar{\mu} I_p(U; Z)] + \bar{\gamma} I_p(X; Y) \} \\
& = \tilde{C}^{(\alpha, \beta, \gamma, \mu)}(W_1, W_2).
\end{aligned}$$

Step (a) follows from (49). Step (b) follows from that by Property 2 part a), the cardinality bound in  $\mathcal{P}(W_1, W_2)$  can be reduced to that in  $\mathcal{P}_{\text{sh}}(W_1, W_2)$ . Step (c) follows from that when  $p \in \mathcal{P}_{\text{sh}}(W_1, W_2)$ , we have

$$D(p_{Y|XZU} \| W_1 | p_{XZU}) = D(p_{Z|XYU} \| W_2 | p_{XYU}) = 0.$$

Hence we have  $\mathcal{C}_{\text{ext}}(W_1, W_2) \subseteq \tilde{\mathcal{C}}_{\text{ext,sh}}(W_1, W_2)$ . Finally we prove  $\tilde{\mathcal{C}}_{\text{ext,sh}}(W_1, W_2) \subseteq \mathcal{C}_{\text{ext,sh}}(W_1, W_2)$ . We assume that  $(\tilde{r}_1, \tilde{r}_2, \tilde{r}_3) \in \tilde{\mathcal{C}}_{\text{ext}}(W_1, W_2)$ . Then we have

$$\begin{aligned}
& \gamma[\mu \tilde{r}_1 + \bar{\mu} \tilde{r}_2] + \bar{\gamma} \tilde{r}_3 \\
& \leq \tilde{C}^{(\alpha, \beta, \gamma, \mu)}(W_1, W_2) \\
& = \max_{q \in \mathcal{Q}} \{ -\alpha D(q_{Y|XZU} \| W_1 | q_{XZU}) \\
& \quad - \beta D(q_{Z|XYU} \| W_2 | q_{XYU}) \\
& \quad + \gamma[\mu I_q(X; Y|U) + \bar{\mu} I_q(U; Z)] \\
& \quad + \bar{\gamma} I_q(X; Y) \} \\
& = -\alpha D(q_{Y|XZU, \alpha, \beta, \gamma, \mu}^* \| W_1 | q_{XZU, \alpha, \beta, \gamma, \mu}^*) \\
& \quad - \beta D(q_{Z|XYU, \alpha, \beta, \gamma, \mu}^* \| W_2 | q_{XYU, \alpha, \beta, \gamma, \mu}^*) \\
& \quad + \gamma[\mu I_{q_{\alpha, \beta, \gamma}^*}(X; Y|U) + \bar{\mu} I_{q_{\alpha, \beta, \gamma}^*}(U; Z)] \\
& \quad + \bar{\gamma} I_{q_{\alpha, \beta, \gamma}^*}(X; Y), \tag{50}
\end{aligned}$$

where  $q_{\alpha, \beta, \gamma, \mu}^* = q_{UXYZ, \alpha, \beta, \gamma, \mu}^* \in \mathcal{Q}$  is a probability distribution which attains the maximum in the definition of  $\tilde{C}^{(\alpha, \beta, \gamma, \mu)}(W_1, W_2)$ . The quantities

$$\begin{aligned}
& q_{Y|XZU, \alpha, \beta, \gamma, \mu}^*, q_{XZU, \alpha, \beta, \gamma, \mu}^*, \\
& q_{Z|XYU, \alpha, \beta, \gamma, \mu}^*, q_{XYU, \alpha, \beta, \gamma, \mu}^*,
\end{aligned}$$

appearing in the right members of (50) are the (conditional) distributions induced by  $q_{\alpha, \beta, \gamma, \mu}^*$ . We set

$$\begin{aligned}
\Delta^{(\gamma, \mu)} & \triangleq \gamma[\mu I_{q_{\alpha, \beta, \gamma, \mu}^*}(X; Y|U) + \bar{\mu} I_{q_{\alpha, \beta, \gamma, \mu}^*}(U; Z)] \\
& \quad + \bar{\gamma} I_{q_{\alpha, \beta, \gamma, \mu}^*}(X; Y) - \gamma[\mu \tilde{r}_1 + \bar{\mu} \tilde{r}_2] - \bar{\gamma} \tilde{r}_3.
\end{aligned}$$

From (50), we must have

$$\begin{aligned}
& 0 \leq \alpha D(q_{Y|XZU, \alpha, \beta, \gamma, \mu}^* \| W_1 | q_{XZU, \alpha, \beta, \gamma, \mu}^*) \\
& \leq \Delta^{(\gamma, \mu)}, \\
& 0 \leq \beta D(q_{Z|XYU, \alpha, \beta, \gamma, \mu}^* \| W_2 | q_{XYU, \alpha, \beta, \gamma, \mu}^*) \\
& \leq \Delta^{(\gamma, \mu)} \tag{51}
\end{aligned}$$

for any  $\alpha, \beta, \gamma, \mu > 0$ . From (51), we have

$$\left. \begin{aligned}
0 & \leq D(q_{Y|XZU, \alpha, \beta, \gamma, \mu}^* \| W_1 | q_{XZU, \alpha, \beta, \gamma, \mu}^*) \\
& \leq \frac{\Delta^{(\gamma, \mu)}}{\alpha}, \\
0 & \leq D(q_{Z|XYU, \alpha, \beta, \gamma, \mu}^* \| W_2 | q_{XYU, \alpha, \beta, \gamma, \mu}^*) \\
& \leq \frac{\Delta^{(\gamma, \mu)}}{\beta}.
\end{aligned} \right\} \tag{52}$$

From (52), we have

$$\begin{aligned}
& \gamma(\mu \tilde{r}_1 + \bar{\mu} \tilde{r}_2) + \bar{\gamma} \tilde{r}_3 \\
& \leq \gamma[\mu I_{q_{\alpha, \beta, \gamma, \mu}^*}(X; Y|U) + \bar{\mu} I_{q_{\alpha, \beta, \gamma, \mu}^*}(U; Z)] \\
& \quad + \bar{\gamma} I_{q_{\alpha, \beta, \gamma, \mu}^*}(X; Y), \tag{53}
\end{aligned}$$

for any  $\alpha, \beta, \gamma$ , and  $\mu > 0$ . Let  $\hat{q}_{\alpha, \beta, \gamma, \mu} = \hat{q}_{UXYZ, \alpha, \beta, \gamma, \mu}$  be a probability distribution with the form

$$\begin{aligned}
& \hat{q}_{UXYZ, \alpha, \beta, \gamma, \mu}(u, x, y, z) \\
& = q_{UX, \alpha, \beta, \gamma, \mu}^*(u, x) W_1(y|x) W_2(z|x).
\end{aligned}$$

Define

$$\begin{aligned}
& \mathcal{Q}(W_1, W_2) \\
& \triangleq \{ q_{UXYZ} : |\mathcal{U}| \leq |\mathcal{Y}| + |\mathcal{Z}| - 1, q_{Y|X} = W_1, q_{Z|X} = W_2, \\
& \quad Y \leftrightarrow X \leftrightarrow Z, U \leftrightarrow X \leftrightarrow (Y, Z) \}.
\end{aligned}$$

By definition, we have  $\hat{q}_{\alpha, \beta, \gamma, \mu} \in \mathcal{Q}(W_1, W_2)$ . Computing  $D(q_{\alpha, \beta, \gamma, \mu}^* \| \hat{q}_{\alpha, \beta, \gamma, \mu})$ , we have the following:

$$\begin{aligned}
& D(q_{\alpha, \beta, \gamma, \mu}^* \| \hat{q}_{\alpha, \beta, \gamma, \mu}) \\
& = D(q_{Y|XZU, \alpha, \beta, \gamma, \mu}^* \| W_1 | q_{XZU, \alpha, \beta, \gamma, \mu}^*) \\
& \quad + D(q_{Z|XYU, \alpha, \beta, \gamma, \mu}^* \| W_2 | q_{XYU, \alpha, \beta, \gamma, \mu}^*) \\
& \quad - I_{q_{\alpha, \beta, \gamma, \mu}^*}(Y; Z|XU) \stackrel{(a)}{\leq} \Delta \left( \frac{1}{\alpha} + \frac{1}{\beta} \right). \tag{54}
\end{aligned}$$

Step (a) follows from (52). From (54), we have

$$\lim_{\alpha, \beta \rightarrow \infty} D(q_{\alpha, \beta, \gamma, \mu}^* \| \hat{q}_{\alpha, \beta, \gamma, \mu}) = 0,$$

from which we have

$$q_{\alpha, \beta, \gamma, \mu}^* \rightarrow \hat{q}_{\alpha, \beta, \gamma, \mu} \text{ as } \alpha, \beta \rightarrow \infty. \tag{55}$$

By (55) and the continuity of  $I_q(X; Y|U)$ ,  $I_q(U; Z)$ , and  $I_q(X; Y)$  with respect to  $q$ , we have that for any  $\gamma, \mu > 0$  and any sufficiently large  $\alpha, \beta$ , we have

$$\begin{aligned}
& \gamma[\mu I_{q_{\alpha, \beta, \gamma, \mu}^*}(X; Y|U) + \bar{\mu} I_{q_{\alpha, \beta, \gamma, \mu}^*}(U; Z)] \\
& \quad + \bar{\gamma} I_{q_{\alpha, \beta, \gamma, \mu}^*}(X; Y) \\
& \leq \gamma[\mu I_{\hat{q}_{\alpha, \beta, \gamma, \mu}}(X; Y|U) + \bar{\mu} I_{\hat{q}_{\alpha, \beta, \gamma, \mu}}(U; Z)] \\
& \quad + \bar{\gamma} I_{\hat{q}_{\alpha, \beta, \gamma, \mu}}(X; Y) + \tau(\alpha, \beta, \gamma, \mu), \tag{56}
\end{aligned}$$

where  $\tau(\alpha, \beta, \gamma, \mu)$  is a positive number that satisfies

$$\lim_{\alpha, \beta \rightarrow \infty} \tau(\alpha, \beta, \gamma, \mu) = 0.$$

Then we have the following chain of inequalities:

$$\begin{aligned}
& \gamma[\mu\tilde{r}_1 + \bar{\mu}\tilde{r}_2] + \bar{\gamma}\tilde{r}_3 \\
\stackrel{(a)}{\leq} & \gamma[\mu I_{q_{\alpha,\beta,\gamma,\mu}^*}(X;Y|U) + \bar{\mu} I_{q_{\alpha,\beta,\gamma,\mu}^*}(U;Z)] \\
& + \bar{\gamma} I_{q_{\alpha,\beta,\gamma,\mu}^*}(X;Y) \\
\stackrel{(b)}{\leq} & \gamma[\mu I_{\hat{q}_{\alpha,\beta,\gamma,\mu}}(X;Y|U) + \bar{\mu} I_{\hat{q}_{\alpha,\beta,\gamma,\mu}}(U;Z)] \\
& + \bar{\gamma} I_{\hat{q}_{\alpha,\beta,\gamma,\mu}}(X;Y) + \tau(\alpha, \beta, \gamma, \mu) \\
\stackrel{(c)}{\leq} & \max_{p \in \mathcal{Q}(W_1, W_2)} \{ \gamma[\mu I_p(X;Y|U) + \bar{\mu} I_p(U;Z)] \\
& + \bar{\gamma} I_p(X;Y) \} + \tau(\alpha, \beta, \gamma, \mu) \\
\stackrel{(d)}{=} & \max_{p \in \mathcal{P}_{\text{sh}}(W_1, W_2)} \{ \gamma[\mu I_p(X;Y|U) + \bar{\mu} I_p(U;Z)] \\
& + \bar{\gamma} I_p(X;Y) \} + \tau(\alpha, \beta, \gamma, \mu) \\
= & C^{(\gamma, \mu)}(W_1, W_2) + \tau(\alpha, \beta, \gamma, \mu). \tag{57}
\end{aligned}$$

Step (a) follows from (53). Step (b) follows from (56). Step (c) follows from that  $\hat{q}_{\alpha,\beta,\gamma,\mu} \in \mathcal{Q}(W_1, W_2)$ . Step (d) follows from that by Property 2 part a), the cardinality  $|\mathcal{U}|$  of  $U$  in  $\mathcal{Q}(W_1, W_2)$  can be upper bounded by  $|\mathcal{X}|$  for describing  $C^{(\gamma, \mu)}(W_1, W_2)$ . Since in (57), the quantity  $\tau(\alpha, \beta, \gamma, \mu)$  can be made arbitrary close to zero, we conclude that  $(\tilde{r}_1, \tilde{r}_2, \tilde{r}_3) \in \mathcal{C}_{\text{ext,sh}}(W_1, W_2)$ . Thus  $\tilde{\mathcal{C}}_{\text{ext,sh}}(W_1, W_2) \subseteq \mathcal{C}_{\text{ext,sh}}(W_1, W_2)$  is proved. ■

### C. Proof of Property 3

In this appendix we prove Property 3.

*Proof of Property 3:* We first prove parts a) and b). For simplicity of notations, set

$$\begin{aligned}
\underline{a} & \triangleq (u, x, y, z), \underline{A} \triangleq (U, X, Y, Z), \underline{\mathcal{A}} \triangleq \mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}, \\
\omega_q^{(\alpha, \beta, \gamma, \mu)}(x, y, z|u) & \triangleq \rho(\underline{a}), \\
\Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2) & \triangleq \xi(\lambda).
\end{aligned}$$

Then we have

$$\Omega^{(\alpha, \beta, \gamma, \mu, \lambda)}(q|W_1, W_2) = \xi(\lambda) = \log \left[ \sum_{\underline{a} \in \underline{\mathcal{A}}} q_{\underline{A}}(\underline{a}) e^{\lambda \rho(\underline{a})} \right].$$

By simple computations we have

$$\begin{aligned}
\xi'(\lambda) &= \left[ \sum_{\underline{a}} q_{\underline{A}}(\underline{a}) e^{\lambda \rho(\underline{a})} \right]^{-1} \left[ \sum_{\underline{a}} q_{\underline{A}}(\underline{a}) \rho(\underline{a}) e^{\lambda \rho(\underline{a})} \right], \tag{58} \\
\xi''(\lambda) &= \left[ \sum_{\underline{a}} q_{\underline{A}}(\underline{a}) e^{\lambda \rho(\underline{a})} \right]^{-2} \\
&\times \left[ \sum_{\underline{a}, \underline{b} \in \underline{\mathcal{A}}} q_{\underline{A}}(\underline{a}) q_{\underline{A}}(\underline{b}) \frac{\{\rho(\underline{a}) - \rho(\underline{b})\}^2}{2} e^{\lambda \{\rho(\underline{a}) + \rho(\underline{b})\}} \right]. \tag{59}
\end{aligned}$$

From (59), it is obvious that  $\xi''(\lambda)$  is nonnegative. Hence  $\Omega(q|W_1, W_2)$  is a convex function of  $\lambda$ . From (58), we have

$$\begin{aligned}
\xi'(0) &= \sum_{\underline{a}} q_{\underline{A}}(\underline{a}) \rho(\underline{a}) \\
&= -\alpha D(q_Y|XZU||W_1|q_{XZU}) \\
&\quad -\beta D(q_Z|XYU||W_2|q_{XYU}) \\
&\quad -\gamma \mu D(q_Y|XU||W_1|q_{XU}) + \gamma \mu I_q(X;Y|U) \\
&\quad + \gamma \bar{\mu} I_q(U;Z) + \bar{\gamma} I_q(X;Y). \tag{60}
\end{aligned}$$

Hence we have the part b). Next we prove the part c). We assume that  $(r_1, r_2, r_3) \notin \mathcal{C}_{\text{ext}}(W_1, W_2)$ , then by Property 2 part b), there exist positive  $\alpha^*, \beta^*, \gamma^*, \mu^*$ , and  $\epsilon$  such that

$$\begin{aligned}
& \gamma^*(\mu^* r_1 + \bar{\mu}^* r_2) + \bar{\gamma}^* r_3 \\
& \geq \tilde{C}^{(\alpha^*, \beta^*, \gamma^*, \mu^*)}(W_1, W_2) + \epsilon. \tag{61}
\end{aligned}$$

Set

$$\begin{aligned}
\zeta(\lambda) &\triangleq \Omega^{(\alpha^*, \beta^*, \gamma^*, \mu^*, \lambda)}(q|W_1, W_2) \\
&\quad - \lambda \left[ -\alpha^* D(q_Y|XZU||W_1|q_{XZU}) \right. \\
&\quad \quad -\beta^* D(q_Z|XYU||W_2|q_{XYU}) \\
&\quad \quad -\gamma^* \mu^* D(q_Y|XU||W_1|q_{XU}) \\
&\quad \quad + \gamma^* [\mu^* I_q(X;Y|U) + \bar{\mu}^* I_q(U;Z)] \\
&\quad \quad \left. + \bar{\gamma}^* I_q(X;Y) + \frac{\epsilon}{2} \right].
\end{aligned}$$

Then we have the following:

$$\left. \begin{aligned} \zeta(0) &= 0, \zeta'(0) = -\frac{\epsilon}{2} < 0, \\ \zeta''(\lambda) &\geq \xi''(\lambda) \geq 0. \end{aligned} \right\} \tag{62}$$

It follows from (62) that there exists  $\nu(\epsilon) > 0$  such that we have  $\zeta(\lambda) \leq 0$  for  $\lambda \in (0, \kappa(\epsilon)]$ . Hence for any  $\lambda \in (0, \nu(\epsilon)]$  and for every  $q \in \mathcal{Q}$ , we have

$$\begin{aligned}
& \Omega^{(\alpha^*, \beta^*, \gamma^*, \mu^*, \lambda)}(q|W_1, W_2) \\
& \leq \lambda \left[ -\alpha^* D(q_Y|XZU||W_1|q_{XZU}) \right. \\
& \quad -\beta^* D(q_Z|XYU||W_2|q_{XYU}) \\
& \quad -\gamma^* \mu^* D(q_Y|XU||W_1|q_{XU}) \\
& \quad + \gamma^* [\mu^* I_q(X;Y|U) + \bar{\mu}^* I_q(U;Z)] \\
& \quad \left. + \bar{\gamma}^* I_q(X;Y) + \frac{\epsilon}{2} \right] \\
& \leq \lambda \left[ -\alpha^* D(q_Y|XZU||W_1|q_{XZU}) \right. \\
& \quad -\beta^* D(q_Z|XYU||W_2|q_{XYU}) \\
& \quad + \gamma^* [\mu^* I_q(X;Y|U) + \bar{\mu}^* I_q(U;Z)] \\
& \quad \left. + \bar{\gamma}^* I_q(X;Y) + \frac{\epsilon}{2} \right]. \tag{63}
\end{aligned}$$

From (63), we have that for any  $\lambda \in (0, \nu(\epsilon)]$ ,

$$\begin{aligned}
& \Omega^{(\alpha^*, \beta^*, \gamma^*, \mu^*, \lambda)}(W_1, W_2) \\
& = \max_{q \in \mathcal{Q}} \Omega^{(\alpha^*, \beta^*, \gamma^*, \mu^*, \lambda)}(q|W_1, W_2) \\
& \leq \lambda \max_{q \in \mathcal{Q}} \left\{ -\alpha^* D(q_Y|XZU||W_1|q_{XZU}) \right. \\
& \quad -\beta^* D(q_Z|XYU||W_2|q_{XYU}) \\
& \quad \left. + \gamma^* [\mu^* I_q(X;Y|U) + \bar{\mu}^* I_q(U;Z)] \right\}
\end{aligned}$$



$$\begin{aligned}
& + \bar{\gamma}^* I_q(X; Y) + \frac{\epsilon}{2} \Big\} \\
& = \lambda \left[ \tilde{C}^{(\alpha^*, \beta^*, \gamma^*, \mu^*)}(W_1, W_2) + \frac{\epsilon}{2} \right]. \tag{64}
\end{aligned}$$

Under (61) and (64), we have the following chain of inequalities:

$$\begin{aligned}
& F_{\text{ext}}(r_1, r_2, r_3 | W_1, W_2) \\
& = \sup_{\substack{\alpha, \beta, \lambda > 0, \\ (\gamma, \mu) \in [0, 1] \\ \times [0, 1/2]}} F_{\text{ext}}^{(\alpha, \beta, \gamma, \mu, \lambda)}(r_1, r_2, r_3 | W_1, W_2) \\
& \geq \sup_{\lambda \in (0, \nu(\epsilon)]} F_{\text{ext}}^{(\alpha^*, \beta^*, \gamma^*, \mu^*, \lambda)}(r_1, r_2, r_3 | W_1, W_2) \\
& = \sup_{\lambda \in (0, \nu(\epsilon)]} \{1 + \lambda[1 + \alpha^* + \beta^* + (2 - 3\mu^*)\gamma^*]\}^{-1} \\
& \quad \times \left[ \lambda \{ \gamma^* (\mu^* r_1 + \bar{\mu}^* r_2) + \bar{\gamma}^* r_3 \} \right. \\
& \quad \left. - \Omega^{(\alpha^*, \beta^*, \gamma^*, \mu^*, \lambda)}(W_1, W_2) \right] \\
& \stackrel{(a)}{\geq} \sup_{\lambda \in (0, \nu(\epsilon)]} \{1 + \lambda[1 + \alpha^* + \beta^* + (2 - 3\mu^*)\gamma^*]\}^{-1} \\
& \quad \times \lambda \left[ \gamma^* (\mu^* r_1 + \bar{\mu}^* r_2) + \bar{\gamma}^* r_3 \right. \\
& \quad \left. - \tilde{C}^{(\alpha^*, \beta^*, \gamma^*, \mu^*)}(W_1, W_2) - \frac{\epsilon}{2} \right] \\
& \stackrel{(b)}{\geq} \sup_{\lambda \in (0, \nu(\epsilon)]} \frac{1}{2} \cdot \frac{\lambda \epsilon}{1 + \lambda[1 + \alpha^* + \beta^* + (2 - 3\mu^*)\gamma^*]} \\
& = \frac{1}{2} \cdot \frac{\nu(\epsilon) \epsilon}{1 + \nu(\epsilon)[1 + \alpha^* + \beta^* + (2 - 3\mu^*)\gamma^*]} > 0.
\end{aligned}$$

Step (a) follows from (64). Step (b) follows from (61). ■

#### D. Proof of Lemma 1

In this appendix we prove Lemma 1. For  $l \in \mathcal{L}_n$ , set

$$\begin{aligned}
& \mathcal{A}_1(l) \\
& \triangleq \left\{ (x^n, y^n, z^n) : \frac{1}{n} \log \frac{p_{Y^n | X^n Z^n L_n}(y^n | x^n, z^n, l)}{q_{Y^n | X^n Z^n L_n}^{(i)}(y^n | x^n, z^n, l)} \right. \\
& \quad \left. = \frac{1}{n} \log \frac{W_1(y^n | x^n)}{q_{Y^n | X^n Z^n L_n}^{(i)}(y^n | x^n, z^n, l)} \geq -\eta \right\}, \\
& \mathcal{A}_2(l) \\
& \triangleq \left\{ (x^n, y^n, z^n) : \frac{1}{n} \log \frac{p_{Z^n | X^n Y^n L_n}(z^n | x^n, y^n, l)}{q_{Z^n | X^n Y^n L_n}^{(ii)}(z^n | x^n, y^n, l)} \right. \\
& \quad \left. = \frac{1}{n} \log \frac{W_2(z^n | x^n)}{q_{Z^n | X^n Y^n L_n}(z^n | x^n, y^n, l)} \geq -\eta \right\}.
\end{aligned}$$

Furthermore, for  $l \in \mathcal{L}_n$ , set

$$\begin{aligned}
& \mathcal{A}_3(l) \triangleq \{(x^n, y^n, z^n) : W_1^n(y^n | x^n) \\
& \quad \geq |\mathcal{K}_n| e^{-n\eta} Q_{Y^n | L_n}^{(iii)}(y^n | l)\}, \\
& \mathcal{A}_4(l) \triangleq \{(x^n, y^n, z^n) : p_{Z^n | L_n}(z^n | l) \\
& \quad \geq |\mathcal{L}_n| e^{-n\eta} Q_{Z^n}^{(iv)}(z^n)\}, \\
& \mathcal{A}_5(l) \triangleq \{(x^n, y^n, z^n) : W_1^n(y^n | x^n) \\
& \quad \geq |\mathcal{K}_n| |\mathcal{L}_n| e^{-n\eta} Q_{Y^n}^{(v)}(y^n)\}, \\
& \mathcal{A}(l) \triangleq \bigcap_{i=1}^5 \mathcal{A}_i(l).
\end{aligned}$$

*Proof of Lemma 1:* We have the following:

$$\begin{aligned}
P_c^{(n)} & = \frac{1}{|\mathcal{K}_n| |\mathcal{L}_n|} \sum_{(k, l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n, z^n) \in \mathcal{A}(l), \\ y^n \in \mathcal{D}_1(k, l), z^n \in \mathcal{D}_2(l)}} \\
& \quad \times \varphi^{(n)}(x^n | k, l) W_1^n(y^n | x^n) W_2^n(z^n | y^n) \\
& \quad + \frac{1}{|\mathcal{K}_n| |\mathcal{L}_n|} \sum_{(k, l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n, z^n) \in \mathcal{A}^c(l): \\ y^n \in \mathcal{D}_1(k, l), z^n \in \mathcal{D}_2(l)}} \\
& \quad \times \varphi^{(n)}(x^n | k, l) W_1^n(y^n | x^n) W_2^n(z^n | y^n) \\
& \leq \sum_{i=0}^5 \Delta_i,
\end{aligned}$$

where

$$\begin{aligned}
\Delta_0 & \triangleq \frac{1}{|\mathcal{K}_n| |\mathcal{L}_n|} \sum_{(k, l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{(x^n, y^n, z^n) \in \mathcal{A}(l)} \\
& \quad \times \varphi^{(n)}(x^n | k, l) W_1^n(y^n | x^n) W_2^n(z^n | x^n), \\
\Delta_i & \triangleq \frac{1}{|\mathcal{K}_n| |\mathcal{L}_n|} \sum_{(k, l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{(x^n, y^n, z^n) \in \mathcal{A}_i^c(l)} \\
& \quad \times \varphi^{(n)}(x^n | k, l) W_1^n(y^n | x^n) W_2^n(z^n | x^n), \\
& \quad \text{for } i = 1, 2, \\
\Delta_i & \triangleq \frac{1}{|\mathcal{K}_n| |\mathcal{L}_n|} \sum_{(k, l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n, z^n) \in \mathcal{A}_i^c(l), \\ y^n \in \mathcal{D}_1(k, l), z^n \in \mathcal{D}_2(l)}} \\
& \quad \times \varphi^{(n)}(x^n | k, l) W_1^n(y^n | x^n) W_2^n(z^n | y^n) \\
& \quad \text{for } i = 3, 4, 5.
\end{aligned}$$

By definition we have

$$\begin{aligned}
& \Delta_0 \\
& = p_{L_n X^n Y^n Z^n} \left\{ \begin{aligned} & 0 \leq \frac{1}{n} \log \frac{W_1^n(Y^n | X^n)}{q_{Y^n | X^n Z^n L_n}(Y^n | X^n, Z^n, L_n)} + \eta, \\ & 0 \leq \frac{1}{n} \log \frac{W_2^n(Z^n | X^n)}{q_{Z^n | X^n Y^n L_n}(Z^n | X^n, Y^n, L_n)} + \eta, \\ & \frac{1}{n} \log |\mathcal{K}_n| \leq \frac{1}{n} \log \frac{W_1^n(Y^n | X^n)}{q_{Y^n | L_n}(Y^n | L_n)} + \eta, \\ & \frac{1}{n} \log |\mathcal{L}_n| \leq \frac{1}{n} \log \frac{p_{Y^n | L_n}(Y^n | L_n)}{q_{Y^n}(Y^n)} + \eta, \\ & \frac{1}{n} \log |\mathcal{L}_n| \leq \frac{1}{n} \log \frac{p_{Z^n | L_n}(Z^n | L_n)}{q_{Z^n}(Z^n)} + \eta \end{aligned} \right\}. \tag{65}
\end{aligned}$$

From (65), it follows that if  $(\varphi^{(n)}, \psi_1^{(n)}, \psi_2^{(n)})$  satisfies

$$\frac{1}{n} \log |\mathcal{K}_n| \geq R_1, \quad \frac{1}{n} \log |\mathcal{L}_n| \geq R_2,$$

then the quantity  $\Delta_0$  is upper bounded by the first term in the right members of (12) in Lemma 1. Hence it suffices to show  $\Delta_i \leq e^{-n\eta}$ ,  $i = 1, 2, 3, 4, 5$  to prove Lemma 1. We first prove  $\Delta_i \leq e^{-n\eta}$  for  $i = 1, 2$ . By a symmetrical structure on  $\mathcal{A}_1(\cdot)$  and  $\mathcal{A}_2(\cdot)$ , it suffices to prove  $\Delta_1 \leq e^{-n\eta}$ . We have the

following chain of inequalities:

$$\begin{aligned}
\Delta_1 &= \sum_{l \in \mathcal{L}_n} \sum_{\substack{(x^n, y^n, z^n) \\ \in \mathcal{A}_1(l)}} p_{L_n X^n Y^n Z^n}(l, x^n, y^n, z^n) \\
&\leq e^{-n\eta} \sum_{l \in \mathcal{L}_n} \sum_{\substack{(x^n, y^n, z^n) \\ \in \mathcal{A}_1(l)}} Q_{Y^n|X^n Z^n L_n}^{(i)}(y^n|x^n, z^n, l) \\
&\quad \times p_{X^n Z^n L_n}(x^n, z^n, l) \\
&\leq e^{-n\eta}.
\end{aligned}$$

Next, we prove  $\Delta_3 \leq e^{-n\eta}$ . We have the following chain of inequalities:

$$\begin{aligned}
\Delta_3 &= \frac{1}{|\mathcal{K}_n||\mathcal{L}_n|} \sum_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n, z^n): \\ y^n \in \mathcal{D}_1(k,l), z^n \in \mathcal{D}_2(l) \\ W_1^n(y^n|x^n) < e^{-n\eta} \\ \times |\mathcal{K}_n| Q_{Y^n|L_n}^{(iii)}(y^n|l)}} \\
&\quad \times \varphi^{(n)}(x^n|k, l) W_1^n(y^n|x^n) W_2^n(z^n|x^n) \\
&\leq \frac{e^{-n\eta}}{|\mathcal{L}_n|} \sum_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n, z^n): \\ y^n \in \mathcal{D}_1(k,l), z^n \in \mathcal{D}_2(l)}} \\
&\quad \times \varphi^{(n)}(x^n|k, l) Q_{Y^n|L_n}^{(iii)}(y^n|l) W_2^n(z^n|x^n) \\
&= \frac{e^{-n\eta}}{|\mathcal{L}_n|} \sum_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{y^n \in \mathcal{D}_1(k,l)} Q_{Y^n|L_n}^{(iii)}(y^n|l) \\
&\quad \times W_2^n(\mathcal{D}_2(l)|x^n) \\
&\leq \frac{e^{-n\eta}}{|\mathcal{L}_n|} \sum_{l \in \mathcal{L}_n} \sum_{k \in \mathcal{K}_n} Q_{Y^n|L_n}^{(iii)}(\mathcal{D}_1(k, l)|l) \\
&= \frac{e^{-n\eta}}{|\mathcal{L}_n|} \sum_{l \in \mathcal{L}_n} Q_{Y^n|L_n}^{(iii)} \left( \bigcup_{k \in \mathcal{K}_n} \mathcal{D}_1(k, l) \middle| l \right) \\
&\leq \frac{e^{-n\eta}}{|\mathcal{L}_n|} \sum_{l \in \mathcal{L}_n} 1 = e^{-n\eta}.
\end{aligned}$$

We prove  $\Delta_4 \leq e^{-n\eta}$ . We have the following chain of inequalities:

$$\begin{aligned}
\Delta_4 &= \frac{1}{|\mathcal{L}_n|} \sum_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n, z^n): \\ y^n \in \mathcal{D}_1(k,l), z^n \in \mathcal{D}_2(l) \\ p_{Z^n|L_n}(z^n|l) < e^{-n\eta} \\ \times |\mathcal{L}_n| Q_{Z^n}^{(iv)}(z^n)}} \\
&\quad \times p_{K_n X^n Y^n Z^n|L_n}(k, x^n, y^n, z^n|l) \\
&\leq \frac{1}{|\mathcal{L}_n|} \sum_{l \in \mathcal{L}_n} \sum_{\substack{z^n \in \mathcal{D}_2(l), \\ p_{Z^n|L_n}(z^n|l) < e^{-n\eta} \\ \times |\mathcal{L}_n| Q_{Z^n}^{(iv)}(z^n)}} \sum_{k \in \mathcal{K}_n} \sum_{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n} \\
&\quad \times p_{K_n X^n Y^n Z^n|L_n}(k, x^n, y^n, z^n|l) \\
&= \frac{1}{|\mathcal{L}_n|} \sum_{l \in \mathcal{L}_n} \sum_{\substack{z^n \in \mathcal{D}_2(l), \\ p_{Z^n|L_n}(z^n|l) < e^{-n\eta} \\ \times |\mathcal{L}_n| Q_{Z^n}^{(iv)}(z^n)}} p_{Z^n|L_n}(z^n|l) \\
&\leq e^{-n\eta} \sum_{l \in \mathcal{L}_n} \sum_{z^n \in \mathcal{D}_2(l)} q_{Z^n}^{(v)}(z^n) = e^{-n\eta} \sum_{l \in \mathcal{L}_n} Q_{Z^n}^{(iv)}(\mathcal{D}_2(l))
\end{aligned}$$

$$= e^{-n\eta} Q_{Z^n}^{(v)} \left( \bigcup_{l \in \mathcal{L}_n} \mathcal{D}_2(l) \right) \leq e^{-n\eta}.$$

Finally, we prove  $\Delta_5 \leq e^{-n\eta}$ . We have the following chain of inequalities:

$$\begin{aligned}
\Delta_5 &= \frac{1}{|\mathcal{K}_n||\mathcal{L}_n|} \sum_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n, z^n): \\ y^n \in \mathcal{D}_1(k,l), z^n \in \mathcal{D}_2(l) \\ W_1^n(y^n|x^n) < e^{-n\eta} \\ \times |\mathcal{K}_n||\mathcal{L}_n| Q_{Y^n}^{(v)}(y^n)}} \\
&\quad \times \varphi^{(n)}(x^n|k, l) W_1^n(y^n|x^n) W_2^n(z^n|x^n) \\
&= \frac{1}{|\mathcal{K}_n||\mathcal{L}_n|} \sum_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n): \\ y^n \in \mathcal{D}_1(k,l), \\ W_1^n(y^n|x^n) < e^{-n\eta} \\ \times |\mathcal{K}_n||\mathcal{L}_n| Q_{Y^n}^{(v)}(y^n)}} \\
&\quad \times \varphi^{(n)}(x^n|k, l) W_1^n(y^n|x^n) W_2^n(\mathcal{D}_2(l)|x^n) \\
&= \frac{1}{|\mathcal{K}_n||\mathcal{L}_n|} \sum_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n): \\ y^n \in \mathcal{D}_1(k,l), \\ W_1^n(y^n|x^n) < e^{-n\eta} \\ \times |\mathcal{K}_n||\mathcal{L}_n| Q_{Y^n}^{(v)}(y^n)}} \\
&\quad \times \varphi^{(n)}(x^n|k, l) W_1^n(y^n|x^n) W_2^n(\mathcal{D}_2(l)|x^n) \\
&\leq \frac{1}{|\mathcal{K}_n||\mathcal{L}_n|} \sum_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \sum_{\substack{(x^n, y^n): \\ y^n \in \mathcal{D}_1(k,l), \\ W_1^n(y^n|x^n) < e^{-n\eta} \\ \times |\mathcal{K}_n||\mathcal{L}_n| Q_{Y^n}^{(v)}(y^n)}} \\
&\quad \times \varphi^{(n)}(x^n|k, l) W_1^n(y^n|x^n) \\
&\leq e^{-n\eta} \sum_{l \in \mathcal{L}_n} \sum_{k \in \mathcal{K}_n} \sum_{\substack{(x^n, y^n): \\ y^n \in \mathcal{D}_1(k,l), \\ W_1^n(y^n|x^n) < e^{-n\eta} \\ |\mathcal{K}_n||\mathcal{L}_n| Q_{Y^n}^{(v)}(y^n)}} \\
&\quad \times \varphi^{(n)}(x^n|k, l) Q_{Y^n}^{(v)}(y^n) \\
&= e^{-n\eta} \sum_{l \in \mathcal{L}_n} \sum_{k \in \mathcal{K}_n} \sum_{\substack{y^n \in \mathcal{D}_1(k,l), \\ W_1^n(y^n|x^n) < e^{-n\eta} \\ |\mathcal{K}_n||\mathcal{L}_n| Q_{Y^n}^{(v)}(y^n)}} Q_{Y^n}^{(v)}(y^n) \\
&\leq e^{-n\eta} \sum_{l \in \mathcal{L}_n} \sum_{k \in \mathcal{K}_n} Q_{Y^n}^{(v)}(\mathcal{D}_1(k, l)) \\
&= e^{-n\eta} Q_{Y^n}^{(v)} \left( \bigcup_{(k,l) \in \mathcal{K}_n \times \mathcal{L}_n} \mathcal{D}_1(k, l) \right) \leq e^{-n\eta}.
\end{aligned}$$

Thus Lemma 1 is proved. ■

#### E. Proof of Lemma 4

In this appendix we prove (27) and (28) in Lemma 4.

*Proofs of (27) and (28) in Lemma 4:* By the definition of  $p_{X^t Y^t|L_n Z^n; \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x^t, y^t|l, z^n)$ , for  $t = 1, 2, \dots, n$ , we have

$$\begin{aligned}
&p_{X^t Y^t|L_n Z^n; \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x^t, y^t|l, z^n) \\
&= C_t^{-1}(l, z^n) p_{X^t Y^t|L_n Z^n; \mathcal{F}^t}(x^t, y^t|l, z^n)
\end{aligned}$$

$$\times \prod_{i=1}^t f_{\mathcal{F}_i}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_i, y_i, z_i | u_i). \quad (66)$$

Then we have the following chain of equalities:

$$\begin{aligned} & p_{X^t Y^t | L_n Z^n; \mathcal{F}^t}(x^t, y^t | l, z^n) \\ \stackrel{(a)}{=} & C_t^{-1}(l, z^n) p_{X^{t-1} Y^{t-1} | L_n Z^n}(x^{t-1}, y^{t-1} | l, z^n) \\ & \times \prod_{i=1}^t f_{\mathcal{F}_i}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_i, y_i, z_i | u_i) \\ = & C_t^{-1}(l, z^n) p_{X^{t-1} Y^{t-1} | L_n Z^n}(x^{t-1}, y^{t-1} | l, z^n) \\ & \times \prod_{i=1}^{t-1} f_{\mathcal{F}_i}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_i, y_i, z_i | u_i) \\ & \times p_{X_t Y_t | X^{t-1} Y^{t-1} L_n Z^n}(x_t, y_t | x^{t-1}, y^{t-1}, l, z^n) \\ & \times f_{\mathcal{F}_t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t | u_t) \\ \stackrel{(b)}{=} & \frac{C_{t-1}(l, z^n)}{C_t(l, z^n)} p_{X^{t-1} Y^{t-1} | L_n Z^n; \mathcal{F}^{t-1}}(x^{t-1}, y^{t-1} | l, z^n) \\ & \times p_{X_t Y_t | X^{t-1} Y^{t-1} L_n Z^n}(x_t, y_t | x^{t-1}, y^{t-1}, l, z^n) \\ & \times f_{\mathcal{F}_t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t, z_t | u_t) \\ = & (\Phi_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n))^{-1} \\ & \times p_{X^{t-1} Y^{t-1} | L_n Z^n; \mathcal{F}^{t-1}}(x^{t-1}, y^{t-1} | l, z^n) \\ & \times p_{X_t Y_t | X^{t-1} Y^{t-1} L_n Z^n}(x_t, y_t | x^{t-1}, y^{t-1}, l, z^n) \\ & \times f_{\mathcal{F}_t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t, z_t | u_t). \quad (67) \end{aligned}$$

Steps (a) and (b) follow from (66). From (67), we have

$$\begin{aligned} & \Phi_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) p_{X^t Y^t | L_n Z^n; \mathcal{F}^t}(x^t, y^t | l, z^n) \quad (68) \\ = & p_{X^{t-1} Y^{t-1} | L_n Z^n; \mathcal{F}^{t-1}}(x^{t-1}, y^{t-1} | l, z^n) \\ & \times p_{X_t Y_t | X^{t-1} Y^{t-1} L_n Z^n}(x_t, y_t | x^{t-1}, y^{t-1}, l, z^n) \\ & \times f_{\mathcal{F}_t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t, z_t | u_t). \quad (69) \end{aligned}$$

Taking summations of (68) and (69) with respect to  $x^t, y^t$ , we obtain

$$\begin{aligned} & \Phi_{t, \mathcal{F}^t}^{(\alpha, \beta, \gamma, \mu, \theta)}(l, z^n) \\ = & \sum_{x^t, y^t} p_{X^{t-1} Y^{t-1} | L_n Z^n; \mathcal{F}^{t-1}}(x^{t-1}, y^{t-1} | l, z^n) \\ & \times p_{X_t Y_t | X^{t-1} Y^{t-1} L_n Z^n}(x_t, y_t | x^{t-1}, y^{t-1}, l, z^n) \\ & \times f_{\mathcal{F}_t}^{(\alpha, \beta, \gamma, \mu, \theta)}(x_t, y_t, z_t | u_t), \end{aligned}$$

completing the proof.  $\blacksquare$

## Acknowledgements

I am very grateful to Dr. Shun Watanabe for his helpful comments.

## REFERENCES

- [1] T. M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-18, no.1, pp. 2-13, Jan. 1972.
- [2] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inform. Theory*, vol. IT-23, no. 1, pp. 60-64, Jan 1977.
- [3] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Academic Press, New York, 1981.
- [4] J. Körner and A. Sgarro, "Universally attainable error exponents for broadcast channels with degraded message sets," *IEEE Trans. Inform. Theory*, vol. IT-26, no. 6, pp. 670-679, Nov. 1980.
- [5] Y. Kaspi and N. Merhav, "Error exponents for broadcast channels with degraded Message sets," *IEEE Trans. Inform. Theory*, vol. 57, no. 1, pp.101-123, Jan. 2011.
- [6] T. S. Han, *Information-Spectrum Methods in Information Theory*. Springer-Verlag, Berlin, New York, 2002. The Japanese edition was published by Baifukan-publisher, Tokyo, 1998.
- [7] Y. Oohama, "Exponent function for one helper source coding problem at rates outside the rate region," *Proceedings of the 2015 IEEE International Symposium on Information Theory*, pp. 1575-1579, Hong Kong, China, June 14-19, 2015.
- [8] Y. Oohama, "Strong converse exponent for degraded broadcast channels at rates outside the capacity region," *Proceedings of the 2015 IEEE International Symposium on Information Theory*, pp. 939 - 943, Hong Kong, China, June 14-19, 2015.
- [9] Y. Oohama, "Strong converse theorems for degraded broadcast channels with feedback," *Proceedings of the 2015 IEEE International Symposium on Information Theory*, pp. 2510 - 2514, Hong Kong, China, June 14-19, 2015.